

2000 Local Similarity 00.00, P-ED: NO. 0.13,
 Matches 44; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 38 ATGCTTTTCGTGCTTGATTAATCTTCTTCGTTTCTCATTAGAATGTAAGCTTCA 97
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 104 ATGCTTCTAAAGTGTGGCATTCTCCGTTCTTCAGGGAATAATTGNAGTCTTGC A 45

Qy 98 AGIT 101
|||
Db 44 AGAT 41

RESULT 2
US-09-010-233-5/C
; Sequence 5, Application US/09010233
; Patent No. 6200774

; GENERAL INFORMATION:
 ; APPLICANT: Black, Michael J.
 ; APPLICANT: Lawlor, Elizabeth J.
 ; APPLICANT: Lewis, Carl J.
 ; TITLE OF INVENTION: NO. 6200774el Compounds
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dechert, Price & Rhoads
 ; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
 ; CITY: Philadelphia
 ; STATE: PA

```

? COUNTRY: USA
? ZIP: 19103-2793
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: Fastseq for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/010.233
?
```

FILING DATE: /
 CLASSIFICATION: /
 PRIOR APPLICATION DATA: /
 APPLICATION NUMBER: 60/037,857 /
 FILING DATE: 07-FEB-1997 /
 APPLICATION NUMBER: 60/044,365 /
 FILING DATE: 28-APR-1997 /
 APPLICATION NUMBER: 60/044,366 /
 FILING DATE: 28-APR-1997 /
 ATTORNEY/AGENT INFORMATION: /
 NAME: Falk, Stephen I /
 REGISTRATION NUMBER: 36,795 /
 REFERENCE/DOCKET NUMBER: GM50027 /
 TELECOMMUNICATION INFORMATION: /
 TELEPHONE: 215-994-2488 /
 TELEFAX: 215-994-2222 /

TELEX: 5
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-010-233-5

```

Query Match      31.7%  Score 32; DB 4; Length 300;
Best Local Similarity 69.8%  Pred No. 0.19;
Matches 44; Conservative 0; Mismatches 20; Indels 0; Gaps

```

QY 38 ATGCCTTTCCGTCGTGGATTAATCTTTCGTCTTCATTAGATAAGTAAAGCTTTCA 97
||||| || | |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 104 ATGCTTCTAAGTGTTGGCAATCTTCGTTCTTCAGGGAATTTGAGCTTGGCA 45

QY 98 AGIT 101
|||
Db 44 AGAT 41

RESULT 3
 US-09-010-232-1/C
 : Sequence 1, Application US/09010232
 : Patent No. 6248557

```

: GENERAL INFORMATION:
: APPLICANT: Black, Michael T.
: APPLICANT: Lawlor, Elizabeth J.
: APPLICANT: Lewis, Ceri J.
: TITLE OF INVENTION: NO. 6248557el Compounds
: NUMBER OF SEQUENCES: 6

```

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dechert, Price & Rhoads
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
 City: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103-3793
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: Fastseq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/010.232
 FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/937,857
 FILING DATE: 07-FEB-1997
 APPLICATION NUMBER: 60/044,366
 FILING DATE: 28-APR-1997
 APPLICATION NUMBER: 60/044,365
 FILING DATE: 28-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T.

REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM50026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

Query Match	31.78;	Score 32;	DB 4;	Length 300;
Best Local Similarity	68.88;			
Matches 4;	Conservative	0;	Mismatches 20;	Indels 0;
				Gaps 0;

QY	38	A T G C C T T T T C C T G C T T G A T A A T T C T G T T C T T C T T C A T T A G A T A T G T A A A T C C T T T C C A	97
Dd	104	A T G C T T C T A A T G T G T G C C A T T C T T C C G T T T C T C A G S A A A A T T T G A A G T C T T C C A	45
QY	98	A G T T	101
Dd	44	A G A T	41

RESULT 4
 US-08-961-527-175/c
 : Sequence 175, Application US/08961527
 : Patent No. 6420135
 : GENERAL INFORMATION:
 : APPLICANT: Charles Kunsch
 : TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 : NUMBER OF SEQUENCES: 391
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Human Genome Sciences, Inc.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 05:54:07 ; Search time 754.132 Seconds
(without alignments)
2086.058 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600

Perfect score: 101

Sequence: 1 atattggtcctgtgtctcc.....tatgtaaatgcttcacagtt 101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

cal number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthma:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_esti:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estium:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pin:*

21: em_gss_vit:*

22: em_gss_fun:*

23: em_gss_nam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
C 1	101	100.0	432	9	AA421341	AA421341 zuc6d11.r
C 2	74.4	73.7	513	9	AI025332	AI025332 CV74903.X
C 3	33.2	32.9	232	17	BH757424	BH757424 SALK_0561
C 4	33.2	32.9	508	17	AQ779483	AQ779483 HS_3001.A
C 5	33	32.7	476	17	A2809047	A2809047 2M0722E9
C 6	32.2	31.9	1197	12	BG538053	BG538053 602563562

C 7	32	31.7	570	13	B1620394	B1620394 603036284
C 8	32	31.7	568	17	BH768691	BH768691 BMBAC361D
C 9	31.6	31.3	403	17	FR0023141	AL106000 F.rubripe
C 10	31.6	31.3	435	9	AL726759	AL726759 AL726759
C 11	31.6	31.3	555	17	AO651385	AO651385 Sheared D
C 12	31.6	31.3	578	9	AL514767	AL514767 AL514767
C 13	31.6	31.3	911	17	AO899744	AO899744 HS_5234..A
C 14	31.4	31.1	534	17	AQ931126	AQ931126 RECI-23-2
C 15	31.2	30.9	452	10	AW988030	AW988030 uf91c01.Y
C 16	31.2	30.9	514	17	BH824692	BH824692 BACP20-1
C 17	31.2	30.9	535	17	A2636018	A2636018 1M0457009
C 18	31.2	30.9	692	17	BH834121	BH834121 BACP66-K1
C 19	31.2	30.9	728	17	BH826207	BH826207 BACP23-E
C 20	31.2	30.9	808	13	B1088928	B1088928 602854293
C 21	31.2	30.9	900	9	AL547213	AL547213 AL547213
C 22	31.2	30.9	1123	12	B3701181	B3701181 602680785
C 23	31	30.7	414	17	A2868761	A2868761 2M0180B09
C 24	31	30.7	636	12	B528560	B528560 602043533
C 25	31	30.7	645	17	A2958131	A2958131 2M0222D02
C 26	31	30.7	881	17	BH152433	BH152433 ENTIP047TR
C 27	31	30.7	941	17	A2532286	A2532286 ENTBM45TF
C 28	31	30.7	1101	17	CNS01501	AL104844 Drosophil
C 29	30.8	30.5	235	9	AI772083	AI772083 EST253183
C 30	30.8	30.5	349	10	AM177332	AM177332 CM1-C1012
C 31	30.8	30.5	512	17	A0765687	A0765687 HS_5372..A
C 32	30.8	30.5	721	12	B3762662	B3762662 602734558
C 33	30.8	30.5	855	17	CNS012BX	AL101367 Drosophil
C 34	30.6	30.3	106	12	B5827565	B5827565 RC2-E1001
C 35	30.6	30.3	482	9	AI317199	AI317199 u135903.Y
C 36	30.6	30.3	534	17	AQ735325	AQ735325 HS_2011..A
C 37	30.6	30.3	583	17	AQ512130	AQ512130 RECI-11-3
C 38	30.6	30.3	596	17	BH259245	BH259245 CH230-123
C 39	30.6	30.3	713	17	A2631576	A2631576 1M0485M18
C 40	30.6	30.3	760	17	A2628465	A2628465 1M0480N05
C 41	30.6	30.3	1067	17	AG029559	AG029559 Pan trogl
C 42	30.4	30.1	309	10	AM177337	AM177337 CM1-C1012
C 43	30.4	30.1	320	10	AM177341	AM177341 CM1-C1012
C 44	30.4	30.1	320	10	AM177381	AM177381 CM4-C1012
C 45	30.4	30.1	321	10	AM177379	AM177379 CM4-C1012

ALIGNMENTS

RESULT 1
AA421341/C
LOCUS
DEFINITION
zuc6d11.r1 Soares_testis_NHr Homo sapiens cDNA clone IMAGE:731061
5' similar to gb:X51602.cdsl VASCULAR ENDOTHELIAL GROWTH FACTOR
RECEPTOR 1 (HUMAN);, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Human.
AA421341.1 GI:2100333
EST.
Human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 432)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gaisel,G., Jost,S.,
Krizmanic,B., Kucab,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tag,F., Theising,B.,
White,I., Wyllie,I., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
COMMENT
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estatson.wustl.edu
This clone is available royalty-free through LINL : contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 801 Std Error: 0.00

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 346.

FEATURES

source

Location/Qualifiers
1. 432
/organism="Homo sapiens"
/db_xref="GDB:5927885"
/db_xref="taxon:9606"
/clone_lib="IMAGE:731061"
/clone_lib="Soares_testis_NHI"
/sex="male"
/lab_host="DH10B"

/note="Vector: p713D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5]. TGTACCAATGAGTGGGCGGCCCAATTTTITTTTITTTT 3'. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p713 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaudo."

SE COUNT

130 a 101 c 93 g 108 t

ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 6.4e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATATGGTCCTGGTCTCTGACATAGAGTGGATGCGCTTTCCTGCTTGATAAT 60

DB 409 ATATGGTCCTGGTCTCTGACATAGAGTGGATGCGCTTTCCTGCTTGATAAT 350

QY 61 TCTTCTGTTTCTTCATTAGATGTAATGCTTTCATGTT 101

DB 349 TCTTCTGTTTCTTCATTAGATGTAATGCTTTCATGTT 309

RESULT 2

AI025332

LOCUS

DEFINITION OV74903.x1 Soares testis.NHI Homo sapiens cDNA clone IMAGE:1643092 3' similar to gb:X51602.cdsl VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 1 (HUMAN); mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 513)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: ccaps-femail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaudo, Ph.D.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbbrp/image/image.html
Insert Length: 841 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 370.

FEATURES

source

Location/Qualifiers
1. 513
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1643092"
/clone_lib="Soares_testis_NHI"

/sex="male"

/lab_host="DH10B"

/note="Vector: p713D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5]. TGTACCAATGAGTGGGCGGCCCAATTTTITTTTITTTT 3'. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p713 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaudo."

BASE COUNT 111 a 129 c 118 g 155 t

ORIGIN

Query Match 73.7%; Score 74.4; DB 9; Length 513;

Best Local Similarity 97.7%; Pred. No. 1.7e-08;

Matches 86; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATATGGTCCTGGTCTCTGACATAGAGTGGATGCGCTTTCCTGCTTGATAAT 60

DB 364 ATATGGTCCTGGTCTCTGACATAGAGTGGATGCGCTTTCCTGCTTGATAAT 423

QY 61 TCTTCTGTTTCTTCATTAGATGTAATGCTTTCATGTT 88

DB 424 TCTTCTGTTTCTTCATTAGATGTAATGCTTTCATGTT 450

RESULT 3

BH757424

LOCUS

DEFINITION BH757424 232 bp DNA linear GSS 01-MAR-2002
SALK_056184.53.00 x Arabidopsis thaliana IDNA insertion lines
Arabidopsis thaliana genomic clone SALK_056184.53.00.x, DNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 232)
Alonso, J.M., Leisner, J.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 3' end of AT5G05190.

Class: TDNA tagged.

FEATURES

source

Location/Qualifiers
1. 232
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone_lib="SALK_056184.53.00.x"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more IDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html."

```

BASE COUNT      73 a   25 c   40 g   79 t   15 others
ORIGIN

Query Match      32.9%; Score 33.2; DB 17; Length 232;
Best Local Similarity 61.6%; Pred. No. 1.7e-02;
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 10 CTGTGCTCTCGACATCAGACGTCGATGCTTTCTCTGCTGTGATGATTCCTTCT 69
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 6 CATGAACTCTCAGAGATCTGAGATGAGAAAGTTTGGCCATGTTTGCCATTCCTTAT 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 70 TCTTCATTAGATATGATTAATGCTTT 95
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 66 CACTTCGTGATAGTAAAGATTT 91
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
A0779483/c
LOCUS      A0779483      508 bp      DNA      linear      GSS 02-AUG-1999
DEFINITION HS.3001.A1.B02.ATC CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-3001 Col-3 Row-C, DNA sequence.
ACCESSION  A0779483
VERSION    A0779483.1 GI:5682443
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE  1 (bases 1 to 508)
AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE      Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE    99380589
COMMENT    Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 3001 row: C column: 3
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 508.
FEATURES
    source
        location/Qualifiers
            1..508
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="Plate-3001 Col-3 Row-C"
                /clone_lib="CIT Approved Human Genomic Sperm Library D"
                /sex="male"
                /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"
BASE COUNT      200 a   93 c   74 g   137 t   4 others
ORIGIN

Query Match      32.9%; Score 33.2; DB 17; Length 508;
Best Local Similarity 61.6%; Pred. No. 1.6e-02;
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 15 GTCTCTGACTCATAGACGTCGATGCTTTCTCTGCTGTGATGATTCCTTCTTCT 74
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 212 GTGCTCTTCTCAICAGATGATGATGATTTCTTCGTGATGATGATTCCTTCTTCT 153
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 75 CATTAGATATGATTAATGCTTTCAAGT 100
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 152 CATTGATTCGTAGGTATTTATAT 127
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 5
A2809047
LOCUS      A2809047      476 bp      DNA      linear      GSS 20-FEB-2001
DEFINITION 2M0072E19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0072E19 R, DNA sequence.
ACCESSION  A2809047
VERSION    A2809047.1 GI:12975012
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 476)
REFERENCE  1
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenan,E., Pedersen,I., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL    Unpublished (2000)
COMMENT    Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: dunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0072 row: E column: 19
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: Plasmid ends
            High quality sequence stop: 476.
FEATURES
    source
        location/Qualifiers
            1..476
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGC2M0072E19"
                /clone_lib="Mouse 10kb plasmid UUGC1M library"
                /sex="Male"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, P-
                /note="Vector: PWD42nv; Purified DNA from M.
                musculus C57BL/6J (male) was obtained from the Jackson
                Laboratory Mouse DNA Resource
                (http://www.jax.org/resources/documents/dnares/). The DNA
                was hydrodynamically sheared by repeated passage through a
                0.005 inch orifice at constant velocity. The sheared DNA
                was blunt end-repaired with T4 DNA polymerase and T4
                polynucleotide kinase. Adaptor oligonucleotides were
                ligated to the blunt ends in high molar excess. The
                adaptor DNA was purified and size-selected for a 9.5 to
                10.5 kb range using preparative agarose gel
                electrophoresis. Vector DNA was prepared from a derivative
                of pWD42 (gil473211+19b/AF129072.1), a copy-number
                inducible derivative of plasmid R1. The vector was ligated
                with adaptors complementary to the insert adaptors and
                purified. The sheared, adaptor mouse DNA was annealed to
                adaptor vector DNA, and transformed into
                chemically-competent E. coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
BASE COUNT      111 a   123 c   100 g   142 t
ORIGIN

Query Match      32.7%; Score 33; DB 17; Length 476;
Best Local Similarity 63.0%; Pred. No. 1.8e-02;
Matches 51; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 17 CTCCTGACCTTAGAGCTGGAGTCCCTTCTCTGCTGTGATGATTCCTTCTTCTTCA 76
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Dd 152 CCCCAGCTGGCAGAGCTTGGTGACGCCCTGCTCTCTCTCTTTTATTTCTCT 211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 77 TTAGATATGTAATGCTTTCA 97
   ||| ||| ||| ||| |||
DB 212 TCCCTATTTCATGATTTTA 232

RESULT 6
B3538053/c
LOCUS B3538053 1197 bp mRNA linear EST 03-APR-2001
DEFINITION 60263362F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:468391 5',
mRNA sequence.
ACCESSION B3538053
VERSION B3538053.1 GI:13530285
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1197)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Clontech Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1502 row: k column: 16
High quality sequence stop: 237.
FEATURES
    source
    location/qualifiers
    1..1197
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:468391"
    /clone_lib="NIH_MGC_77"
    /lab_host="DH10B (T1 phage-resistant)"
    /note="Organ: lung; Vector: pDR-LIB (Clontech); Site: 1:
    Sfil (ggcgctggcc); Site 2: Sfil (ggcattatggcc); 5' and
    3' adaptors were used in cloning as follows: 5' adaptor
    sequence: 5'-CAGGCCATTATGCCC-3' and 3' adaptor sequence:
    5'-ATTATGAGCGGAGCGGCCGACATG-dt(30)BN-3' (where B = A,
    C, or G and N = A, C, G, or T). Average insert size 1.9
    kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
    by PCR. This library was enriched for full-length clones
    and was constructed by Clontech Laboratories (Palo Alto,
    CA). Note: this is a NIH_MGC Library."
E COUNT 512 a 284 c 246 g 155 t
.IGIN

Query Match 31.9%; Score 32.2; DB 12; Length 1197;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 46; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 28 TTAGCTGATGCGCTTTCCTGCTTTGTAATTTCTTCTTCATTAGATATGA 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 842 TTCTGTGATGTGTGCTCCCTGCTTTTCTGCTTCTTCTTCTTCTTCTTCTT 783
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 88 AATGCTTTC 96
   ||| |||
DB 782 CCTCGCTC 774

RESULT 7
B1820394/c
LOCUS B1820394 570 bp mRNA linear EST 04-OCT-2001
DEFINITION 603036284F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5177156 5',
mRNA sequence.
ACCESSION B1820394
VERSION B1820394.1 GI:15931944
KEYWORDS EST.

```

```

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 570)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11441 row: h column: 21
High quality sequence stop: 451.
FEATURES
    source
    location/qualifiers
    1..570
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5177156"
    /clone_lib="NIH_MGC_115"
    /lab_host="DH10B"
    /note="Organ: pooled brain, lung, testis; Vector:
    PCMV-SPOR16; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA
    source anonymous pool of 6 male brains, age range 23-27; 1
    male lung, age 27; and 1 male testis, age 69. Library is
    oligo-dT primed and directionally cloned (EcoRV site is
    destroyed upon cloning). Average insert size 1.8 kb,
    insert size range 1-3 kb. Library is normalized and
    enriched for full-length clones and was constructed by C.
    Gruber (Invitrogen). Research Genetics tracking code
    021. Note: this is a NIH_MGC Library."
BASE COUNT 158 a 138 c 169 g 105 t
ORIGIN

Query Match 31.7%; Score 32; DB 13; Length 570;
Best Local Similarity 62.5%; Pred. No. 3.1e+02;
Matches 50; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 10 CTGTGCTTCCTGCTCATTTAGACCTGATGCGCTTCTGCTGTTGATATCTTCTGT 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 508 CIGGCATCTTCGCTGCTTTATCTTCGCTTCGCTTCATCTTCATCTTCATCTTC 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 TCTTCATTAGATATGTA 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 448 TTTTTCATTAGATGTA 429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
B1826691/c
LOCUS B1826691 668 bp DNA linear GSS 20-MAR-2002
DEFINITION BMBAC361D117_P5U Brugia malayi Genomic Bac Library 3 Brugia malayi
genomic, DNA sequence.
ACCESSION B1826691
VERSION B1826691.1 GI:19566455
KEYWORDS GSS.
SOURCE Brugia malayi.
ORGANISM Brugia malayi
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Brugia.
REFERENCE 1 (bases 1 to 668)
AUTHORS Whitton, C., Daub, J., Ware, J., Quail, M., Hall, N., Barrall, B., Foster
J., Guiliano, D., Slatko, B. and Blaxter, M.
TITLE Genome survey sequences from the human parasitic nematode Brugia
malayi
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter, M.L.
Institute of Cell, Animal and Population Biology
University of Edinburgh

```


REFERENCE
 A0651385.1 GI:5144571
 GSS.
 Trypanosoma brucei.
KEYWORDS
 Trypanosoma brucei
ORGANISM
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.
REFERENCE
 1 (bases 1 to 555)
 El-Sayed, M., Zhao, S., Zhao, H., Gill, S., Suh, E., Malek, J., Fujii, C.,
 Gerard, C., Leech, V., de Jong, P., Ullu, E., Melville, S., Donelson, J.,
 Fraser, C., and Adams, M.
TITLE
 Determination of clone end sequences from Trypanosoma brucei GUTat
 10.1 sheared DNA library
JOURNAL
 Other_GSSs: Sheared DNA-17L3.TF
COMMENT
 Contact: Najib M. El-Sayed
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: nelsayed@tigr.org
 Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
 DNA library constructed at TIGR. Clones will be available for
 distribution through ATCC. Sheared DNA end sequences search page:
<http://www.tigr.org/tcb/mdb/tbdb/>
 Seq primer: M13-Reverse
 Class: shotgun.
FEATURES
 Location/Qualifiers
 source
 1..555
 /organism="Trypanosoma brucei"
 /strain="IREU927/4 GUTat 10.1"
 /db_xref="taxon:5691"
 /clone="Sheared DNA-17L3"
 /clone_lib="Sheared DNA"
 /note="Vector: pUC18; Site 1: SmaI; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Trypanosoma brucei isolated from a cloned population of
 Trypanosoma brucei (IREU927/4 GUTat 10.1) was mechanically
 sheared to give a tight size distribution (approx 2 kb).
 The v + i method used for the library construction is
 described in detail in Smith, H.O. and Venter, J.C.
 (Making small insert libraries for whole genome shotgun
 sequencing projects. In Genome Sequencing: A Practical
 Approach, eds. M. Vaubin and B. Sorell, Oxford University
 Press, 1995)."
 BASE COUNT 207 a 83 c 85 g 180 t
 ORIGIN
 Query Match 31.3%; Score 31.6; DB 17; Length 555;
 est Local Similarity 62.8; Pred. No. 3.8e+02;
 Matches 49; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 18 TCCTGACTATTAGAGCTGGATGCTTTCTCTGCTGATTAATCTTCTGTTCTCAT 77
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 DB 465 TCCTCTGCTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 524
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 78 TAGATATGTAATGCTTT 95
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 DB 525 TAAATTTATGAATGCTTT 542
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
RESULT 12
 AL514767/c
LOCUS
 DEFINITION AL514767 LTR_NFL006.P12 Homo sapiens cDNA clone CLOB0152602 3
 prime, mRNA sequence.
 ACCESSION AL514767
 VERSION AL514767
 KEYWORDS EST.
SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1 (bases 1 to 578)
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
JOURNAL
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
 Location/Qualifiers
 source
 1..578
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CLOB0152602"
 /clone_lib="LTR_NFL006.P12"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9600 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com>"
 BASE COUNT 150 a 119 c 135 g 138 t 36 others
 ORIGIN
 Query Match 31.3%; Score 31.6; DB 9; Length 578;
 Best Local Similarity 58.1%; Pred. No. 3.8e+02;
 Matches 43; Conservative 6; Mismatches 25; Indels 0; Gaps 0;
 QY 17 CTCCTGACTATTAGAGCTGGATGCTTTCTCTGCTGATTAATCTTCTGTTCTTCA 76
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 DB 150 CTCCTGACTATTAGAGCTGGATGCTTTCTCTGCTGATTAATCTTCTGTTCTTCA 91
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 77 TTGATAATGTAAT 90
 || || || || || || || || || || || || || || || || || || || || ||
 DB 90 TTTTCTTTGTAAT 77
RESULT 13
 A0895744
LOCUS
 DEFINITION HS-5234.AL.F10.T7A RPI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=5002 Col=19 Row=K, DNA sequence.
 ACCESSION A0895744
 VERSION A0895744.1 GI:6355934
 KEYWORDS GSS.
SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1 (bases 1 to 911)
AUTHORS
 Mahairs, G.G., Wallace, J.C., Smith, K., Searzeil, S., Holzman, T.,
 Keller, A., Shaker, A., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
 Hood, L.
TITLE
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
JOURNAL
 MEDLINE
COMMENT
 Contact: Mahairs GC, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98105, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BAC/RAC Resources (<http://bacpac.med.buffalo.edu/ordering.bac.htm>)
 or from Resear h Genetics (info@resgen.com). BAC end Web Server:

Matches 51; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2 TATGTCCTGTGCTCTCCCTGACTCATAGAGCTGGATGCCCTTTCCCTGCTGATTAAT 61

DB 339 TTCTCTGATTTGACGATCAGAGTTTTCACCTGCTGGCTGCATTTTCTCTCTGCAAT 280

QY 62 CTTCCTGTTTCTTCATTAGATATGT 86

DB 279 CTTTGTGTACACCACTAATAAT 255

Search completed: December 8, 2002, 07:50:31
For time : 790.132 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:10:23 ; Search time 26.618 Seconds
(without alignments)
1478.802 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600

Perfect score: 101
Sequence: 1 atgttgctgtgtctcc.....tatgtaaatgttcaagtt 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 350425 seqs, 194966369 residues

cal number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match %	Length	DB ID	Description				
C 1	32	31.7	288	10	US-09-815-242-4447	Sequence 4117, Ap			
C 2	32	31.7	303	10	US-09-815-242-8393	Sequence 8393, Ap			
C 3	32	31.7	303	10	US-09-815-242-8767	Sequence 8767, Ap			
C 4	32	31.7	303	10	US-09-815-242-8943	Sequence 8943, Ap			
C 5	30.6	30.3	3931	12	US-10-051-952-5	Sequence 5, Appli			
C 6	29.2	28.9	429	9	US-10-001-887-73	Sequence 73, Appl			
C 7	28.6	28.3	75899	10	US-09-854-883-243	Sequence 243, App			
C 8	28	27.7	5000	10	US-09-864-761-10713	Sequence 10713, A			
C 9	28	27.7	5000	10	US-09-791-105-2	Sequence 2, Appli			
C 10	28	27.7	76798	10	US-09-880-107-3949	Sequence 3949, Ap			
C 11	28	27.7	397658	10	US-09-813-320-3	Sequence 3, Appli			
C 12	27.8	27.5	4797	10	US-09-751-797-25	Sequence 25, Appl			
C 13	27.4	27.1	397658	10	US-09-813-320-3	Sequence 3, Appli			
C 14	27.2	26.9	21045	10	US-09-764-864-1695	Sequence 1695, Ap			
C 15	27	26.7	147309	10	US-09-742-312-3	Sequence 3, Appli			
C 16	27	26.7	1503841	9	US-09-946-807-1	Sequence 1, Appli			
C 17	27	26.7	1503841	10	US-09-795-668-1	Sequence 1, Appli			
C 18	27	26.7	1503841	10	US-09-795-668-1	Sequence 1, Appli			
C 19	26.8	26.5	432	10	US-09-864-761-33271	Sequence 33271, A			

20	26.8	26.5	555	10	US-09-864-761-16805	Sequence 16805, A
21	26.8	26.5	24533	9	US-09-764-868-1349	Sequence 1349, Ap
C 22	26.6	26.3	387	10	US-09-864-761-10466	Sequence 10466, A
C 23	26.6	26.3	2000	9	US-09-938-842A-5172	Sequence 5172, Ap
C 24	26.6	26.3	2529	10	US-09-764-847-1859	Sequence 1859, Ap
C 25	26.6	26.3	17569	9	US-09-736-457-1804	Sequence 1804, Ap
C 26	26.6	26.3	17569	9	US-09-902-941-1804	Sequence 1804, Ap
C 27	26.6	26.3	115592	10	US-09-818-512-3	Sequence 3, Appli
C 28	26.4	26.1	1805	10	US-09-728-422-7	Sequence 7, Appli
C 29	26.4	26.1	2724	10	US-09-511-781-31	Sequence 31, Appli
C 30	26.4	26.1	3164	10	US-09-764-869-1451	Sequence 1451, Ap
C 31	26.4	26.1	3571	9	US-10-135-583-42	Sequence 42, Appli
C 32	26.4	26.1	3571	10	US-09-823-033-3	Sequence 3, Appli
C 33	26.4	26.1	4500	10	US-09-729-653-1	Sequence 1, Appli
C 34	26.4	26.1	14171	10	US-09-764-877-3517	Sequence 3517, Ap
C 35	26.4	26.1	15958	10	US-09-764-877-3518	Sequence 3518, Ap
C 36	26.2	25.9	2309	10	US-09-925-300-680	Sequence 680, App
C 37	26.2	25.9	2457	10	US-09-739-254-40	Sequence 40, Appli
C 38	26.2	25.9	2457	10	US-09-904-615-40	Sequence 40, Appli
C 39	26.2	25.9	38374	10	US-09-880-107-3463	Sequence 3463, Ap
C 40	26.2	25.9	249467	9	US-10-026-188-3	Sequence 3, Appli
C 41	26	25.7	347	10	US-09-783-590-7481	Sequence 7481, Ap
C 42	26	25.7	450	10	US-09-834-975-622	Sequence 622, App
C 43	26	25.7	450	10	US-09-834-975-763	Sequence 763, App
C 44	26	25.7	656	10	US-09-878-574-4414	Sequence 4414, Ap
C 45	26	25.7	1381	10	US-09-925-301-276	Sequence 276, App

ALIGNMENTS

RESULT 1
US-09-815-242-4447/C.
; Sequence 4447, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Irawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ. ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4447
; LENGTH: 288
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4447

Query Match 31.7% Score 32; DB 10; Length 288;
Best Local Similarity 66.8%; Pred. No. 2.2;

DB 1656 TGAT 1659

RESULT 15
US-09-742-312-3/c
; Sequence 3; Application US/09742312
; Patent No. US20020045166A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000838
; CURRENT APPLICATION NUMBER: US/09/742,312
; CURRENT FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 147309
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(147309)
; OTHER INFORMATION: n = A,T,C or G
US-09-742-312-3

Query Match 26.7%; Score 27; DB 10; Length 147309;
Best Local Similarity 66.1%; Pred. No. 1.8e+02;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 23 ACTCATTAGAGCTGGATGCTTTTCCTGCTGATAATCTTCTGTTCTTCATTAGA 81
DB 49881 AATCATTAGTGTAGTGTCCAAATAGCCCGGATTCCTTCGGCTCTGTGTA 49823

Search completed: December 8, 2002, 09:29:51
Job time : 274.632 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuDen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:52:47 : Search time 5277.36 Seconds
(without alignments)
11586.276 Million cell updates/sec

Title: X51602_COPY_1900_4000
Perfect score: 2101
Sequence: 1 TATATACAGAGATGCGCAA.....CCACAGCAGCAGCTGCTGCG 2101

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
4109280
total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank:	1: gb-be.*	X51602 Human flt m
2: gb-hgt.*	2: 2097.8	AX481481 Sequence
3: gb-in.*	3: 2097.8	AF063657 Homo sapi
4: gb-om.*	4: 1599.2	D28498 Rat mRNA fo
5: gb-ov.*	5: 1574.5	L07297 Mus musculu
6: gb-pat.*	6: 1574.5	D88689 Mus musculu
7: gb-ph.*	7: 1534.2	X78568 M.musculus
8: gb-pl.*	8: 1168.8	AB065372 Gallus ga
9: gb-pr.*	9: 956.4	AB065372 Gallus ga
10: gb-ro.*	10: 953.2	AB065372 Gallus ga
11: gb-sts.*	11: 953.2	AB065372 Gallus ga
12: gb-sy.*	12: 953.2	AB065372 Gallus ga
13: gb-un.*	13: 953.2	AB065372 Gallus ga
14: gb-vi.*	14: 953.2	AB065372 Gallus ga
15: em-ba.*	15: 953.2	AB065372 Gallus ga
16: em-fun.*	16: 953.2	AB065372 Gallus ga
17: em-hum.*	17: 953.2	AB065372 Gallus ga
18: em-in.*	18: 953.2	AB065372 Gallus ga
19: em-mu.*	19: 953.2	AB065372 Gallus ga
20: em-om.*	20: 953.2	AB065372 Gallus ga
21: em-or.*	21: 953.2	AB065372 Gallus ga
22: em-ov.*	22: 953.2	AB065372 Gallus ga
23: em-ph.*	23: 953.2	AB065372 Gallus ga
24: em-pl.*	24: 953.2	AB065372 Gallus ga
25: em-fo.*	25: 953.2	AB065372 Gallus ga
26: em-fo.*	26: 953.2	AB065372 Gallus ga
27: em-sts.*	27: 953.2	AB065372 Gallus ga
28: em-un.*	28: 953.2	AB065372 Gallus ga
29: em-vi.*	29: 953.2	AB065372 Gallus ga
30: em-hgt-hum.*	30: 953.2	AB065372 Gallus ga
31: em-hgt-in.*	31: 953.2	AB065372 Gallus ga
32: em-hgt-other.*	32: 953.2	AB065372 Gallus ga
33: em-hgt-mus.*	33: 953.2	AB065372 Gallus ga
34: em-hgt-pln.*	34: 953.2	AB065372 Gallus ga
35: em-hgt-rod.*	35: 953.2	AB065372 Gallus ga
36: em-hgt-man.*	36: 953.2	AB065372 Gallus ga
37: em-hgt-vrt.*	37: 953.2	AB065372 Gallus ga
38: em-sf.*	38: 953.2	AB065372 Gallus ga
39: em-hgtgo-hum.*	39: 953.2	AB065372 Gallus ga
40: em-hgtgo-mus.*	40: 953.2	AB065372 Gallus ga
41: em-hgtgo-other.*	41: 953.2	AB065372 Gallus ga

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2101	100.0	7680	9	HSFLT
2	2097.8	99.8	4017	6	AX481481
3	2097.8	99.8	4017	9	AF063657
4	1599.2	76.1	4734	10	RAIFL1
5	1574.5	74.9	6055	10	MUSRTKB
6	1574.5	74.9	6055	10	D88689
7	1534.2	73.0	5482	10	MMEMRK2
8	1168.8	55.6	4272	5	AB065372
9	956.4	45.5	22065	2	HSN12688
10	953.2	45.4	197279	9	AC104439
11	953.2	45.4	197279	9	AC104439
12	713.8	34.0	4044	5	AX481480
13	713.8	34.0	4071	6	AR142803
14	713.8	34.0	4071	6	AR201382
15	713.8	34.0	4071	6	AX481483
16	713.8	34.0	4225	6	AF063658
17	713.8	34.0	5830	9	AF035121
18	710.6	33.8	4230	9	HSGRFK
19	710.6	33.8	4230	9	HUNKDRZ
20	710.6	33.8	4236	5	AR030758
21	710.6	33.8	5464	10	BC020530
22	710.6	33.8	5470	6	AR068047
23	710.6	33.8	5470	6	AR125598
24	710.6	33.8	5470	6	AR068047
25	710.6	33.8	5892	10	MMFLK1
26	708.4	33.7	5391	10	MMFLK1
27	708.4	33.7	5406	6	AR005213
28	708.4	33.7	5406	6	AR071705
29	708.4	33.7	5406	6	125171
30	708.4	33.7	5406	6	140602
31	705.8	33.6	4790	10	S53103
32	690	33.0	5390	5	CCQUEK1
33	690	32.8	2352	6	AR030849
34	690	32.8	2352	6	182808
35	665.8	31.7	5037	5	AF180354
36	661	31.5	3906	5	AF056466
37	640.6	30.5	4617	5	CCQUEK2
38	624	29.7	2523	6	AR163540
39	624	29.7	2523	6	E13256
40	609.6	29.0	3009	6	E14000
41	551.8	26.7	5284	10	MUSRTKA
42	541.6	25.8	3897	6	AX481482
43	541.6	25.8	4195	6	AR016568
44	541.6	25.8	4195	6	AR106405
45	541.6	25.8	4195	6	AR106405

ALIGNMENTS

RESULT 1
HSFLT Human flt mRNA for receptor-related tyrosine kinase.
LOCUS X51602
DEFINITION X51602.1 GI:31431
ACCESSION X51602
VERSION 1
KEYWORDS flt gene; fms-related tyrosine kinase gene; tyrosine kinase.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 7680)
AUTHORS Fukuyama; Metaxas; Chordata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL Direct Submission
Submitted (02-JAN-1989) Shibuya M., Institute of Medical Science,
PRI 15-NOV-1993

University of Tokyo, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108.
 Japan
 2 (bases 1 to 7680)
 Shibuya M., Yamaguchi S., Yamane A., Ikeda I., Tojo A.,
 Matsushita H. and Sato M.
 Nucleotide sequence and expression of a novel human receptor-type
 tyrosine kinase gene (flt) closely related to the fms family
 Oncogene 5 (4), 519-524 (1990)
 90221591
 PUBMED
 2158038
 3 (bases 1 to 7680)
 Han H.J., Fujiwara I., Shin S. and Nakamura Y.
 Dinucleotide repeat polymorphism in the 3' non-coding region of the
 flt1 gene
 Hum. Mol. Genet. 2 (12), 2204 (1993)
 94154724
 PUBMED
 8111406
 Data kindly reviewed (20-JUL-1990) by Shibuya M.
 Location/Qualifiers
 1. 7680
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="13"
 /map="13q12"
 /clone="3-7, 3-5"
 /tissue_type="placenta"
 250...4256
 /note="flt gene product (AA 1-1338)"
 /codon_start=1
 /protein_id="CAA35946.1"
 /db_xref="GI:31432"
 /translation="MVSYWDIVGLLCALLSCILLTGSSSGSKLDPFLSLKGTQHIWQ
 AGTLLHLCRGAAHKKWLEPMGSKESRLSIKACNGKNGKFCSTLILNKTAQNHQ
 GFYSYALVPTSKKTEKATFISDTRFVEMYEIPELIMTEGPELVIPCRV
 ISNIVILKPELIDIPGKRIINDSKGEIISNATYKELIGLICEATYKSHLYKI
 SVLRHONTLIDVOISTPRVLLRHLVNLCTAIPLNIRQMTWISYIPDKRKA
 TVKRRQVLEIVAGKRSYRLSKMKVAFSPVWLKDLGPAEKARYLIRYSLI
 KVIDEAGNTILLISUKOSWFKNLATILVNVKPOIYKAVSFDPDAIPYLGSR
 ILITAYGIPQPTIKWFWHPCHNHNSEARCFDCSNNEESFLDANSNMGNLESTIQ
 MAIEGKWKASLIVADSPISGIICTIASNKVIGVNRNLSFVIDVNGFVNLKRW
 PTEGDLKLCSTNKKLYRDVITWILLRVNRIWYISYKQMAIKHESILNLTIV
 NVLSQDSGTCARVYTGEEILKKEIIRDEAPYLRLNSDHIWAISSITLDC
 HANGVPEPQITWFKNNHKIQOEPGIIIGPSSILFHERVTEDEGVHCKATNKGW
 ESSALYVOGTSKSNLELITICTVAAILFLLWLLTLIRKRSSSFIAKIDVLSI
 MDPEVLPDQERLPYDASKWEFARLKLKSLGRGAFKVVQASAFGLKAFICP
 TVAVMLKEGATASEYKALMELKILTHIGHLNVNLLCACTKOGPLKLVYKCYK
 EDKLSLSPVEEDSUGFYKEPITMEDLISYFOVARGMEFLSSRCKLRDLAENLL
 SENNVYKICDGLARDIYNPDYVRKGDITLPIKWNAPESIFDKIYSTKSYRSGVL
 LAELFSLGSPYGVOMDEDFCSRLREGMRAPESYPIYQIMLDCWHRDKRPPR
 FAEVLKGLDQANVOODKQYIPINAILLGNSTYTSYFASDFKESISAPFN
 SGSDVRYVNAFKWMLERIKTEFELLNPMEDDYQDSSTILLASPVLRKEIWD
 SKFKSLKLDLAVISKSEGLSDVSRPSCSHSGHVSSEKGRFTVDHAELEPKIAC
 CSPPEYNSVLYSTPPI"
 2279 a 1561 c 1739 g 2001 t

CDS

1. 7680
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="13"
 /map="13q12"
 /clone="3-7, 3-5"
 /tissue_type="placenta"
 250...4256
 /note="flt gene product (AA 1-1338)"
 /codon_start=1
 /protein_id="CAA35946.1"
 /db_xref="GI:31432"
 /translation="MVSYWDIVGLLCALLSCILLTGSSSGSKLDPFLSLKGTQHIWQ
 AGTLLHLCRGAAHKKWLEPMGSKESRLSIKACNGKNGKFCSTLILNKTAQNHQ
 GFYSYALVPTSKKTEKATFISDTRFVEMYEIPELIMTEGPELVIPCRV
 ISNIVILKPELIDIPGKRIINDSKGEIISNATYKELIGLICEATYKSHLYKI
 SVLRHONTLIDVOISTPRVLLRHLVNLCTAIPLNIRQMTWISYIPDKRKA
 TVKRRQVLEIVAGKRSYRLSKMKVAFSPVWLKDLGPAEKARYLIRYSLI
 KVIDEAGNTILLISUKOSWFKNLATILVNVKPOIYKAVSFDPDAIPYLGSR
 ILITAYGIPQPTIKWFWHPCHNHNSEARCFDCSNNEESFLDANSNMGNLESTIQ
 MAIEGKWKASLIVADSPISGIICTIASNKVIGVNRNLSFVIDVNGFVNLKRW
 PTEGDLKLCSTNKKLYRDVITWILLRVNRIWYISYKQMAIKHESILNLTIV
 NVLSQDSGTCARVYTGEEILKKEIIRDEAPYLRLNSDHIWAISSITLDC
 HANGVPEPQITWFKNNHKIQOEPGIIIGPSSILFHERVTEDEGVHCKATNKGW
 ESSALYVOGTSKSNLELITICTVAAILFLLWLLTLIRKRSSSFIAKIDVLSI
 MDPEVLPDQERLPYDASKWEFARLKLKSLGRGAFKVVQASAFGLKAFICP
 TVAVMLKEGATASEYKALMELKILTHIGHLNVNLLCACTKOGPLKLVYKCYK
 EDKLSLSPVEEDSUGFYKEPITMEDLISYFOVARGMEFLSSRCKLRDLAENLL
 SENNVYKICDGLARDIYNPDYVRKGDITLPIKWNAPESIFDKIYSTKSYRSGVL
 LAELFSLGSPYGVOMDEDFCSRLREGMRAPESYPIYQIMLDCWHRDKRPPR
 FAEVLKGLDQANVOODKQYIPINAILLGNSTYTSYFASDFKESISAPFN
 SGSDVRYVNAFKWMLERIKTEFELLNPMEDDYQDSSTILLASPVLRKEIWD
 SKFKSLKLDLAVISKSEGLSDVSRPSCSHSGHVSSEKGRFTVDHAELEPKIAC
 CSPPEYNSVLYSTPPI"

BASE COUNT

2279 a 1561 c 1739 g 2001 t

Query Match

Best Local Similarity 100.0%; Score 2101; DB 9; Length 7680;

Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATATCAGAGTGTGCAATGGTTCATGTTAACTTGGAAATAATCCGACGAGGA 60
 DB 1900 TATATCAGAGTGTGCAATGGTTCATGTTAACTTGGAAATAATCCGACGAGGA 60
 QY 61 GAGGACCTGAACCTGCTTCACAGCTTACAGTCTTATACAGAGCTTACTTGAAT 120
 DB 1960 GAGGACCTGAACCTGCTTCACAGCTTACAGTCTTATACAGAGCTTACTTGAAT 120
 QY 121 TTACTGGGACAGTTAATACAGAACAAATGACATGATATAGCAAGCAAAAATGGCC 160
 DB 3100 CAAGGCAAGAACCAAGACTAGATAGCTCNCACGACGGAAGCTTTGGAGCTCCGCG 3159

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY	1261	TTTCAGGAAGATATAAGATCTGAGTCGATGTTGAGAGAGAGAGATCTGACGGTTCTAC	1320
Db	3160	TTTCAGGAAGATATAAGATCTGAGTCGATGTTGAGAGAGAGAGATCTGACGGTTCTAC	3219
QY	1321	AAGGAGCCCATCTACTATGGAAGATCTGATTTCTTACAGTTTTCAAGTGGTCAAGAGCATG	1380
Db	3220	AAGGAGCCCATCTACTATGGAAGATCTGATTTCTTACAGTTTTCAAGTGGTCAAGAGCATG	3279
QY	1381	GAGTTCTGTCTCTCAGAAAGTGGATCATCGGAGCTGGCAGCCAGCAAAATCTTTTA	1440
Db	3280	GAGTTCTGTCTCTCAGAAAGTGGATCATCGGAGCTGGCAGCCAGCAAAATCTTTTA	3339
QY	1441	TCTGAGAACACGCTGGTGAAGATTTGTGATTTGGCTTCCCGGATATTTATAAGAAC	1500
Db	3340	TCTGAGAACACGCTGGTGAAGATTTGTGATTTGGCTTCCCGGATATTTATAAGAAC	3399
QY	1501	CCGGATTAATGTGAGAAAGAGATCTCTCGAGTTCTCTCTGAAATGGATGGTCCCGAATCT	1560
Db	3400	CCGGATTAATGTGAGAAAGAGATCTCTCGAGTTCTCTCTGAAATGGATGGTCCCGAATCT	3459
QY	1561	ATCTTTGACAAATCTACAGCAACAGAGCGATGTGTCTTACCGAGATTCTCTGCG	1620
Db	3460	ATCTTTGACAAATCTACAGCAACAGAGCGATGTGTCTTACCGAGATTCTCTGCG	3519
QY	1621	GAATCTTCTCTTAGTGGTCTCCATACCCAGGAGTACAAATGGATGAGGACTTTTC	1680
Db	3520	GAATCTTCTCTTAGTGGTCTCCATACCCAGGAGTACAAATGGATGAGGACTTTTC	3579
QY	1681	AGTCGCTTGAGGAGGCGATGAGGATGAGAGCTCTGAGTACTTACTCTTGAAATCTAT	1740
Db	3580	AGTCGCTTGAGGAGGCGATGAGGATGAGAGCTCTGAGTACTTACTCTTGAAATCTAT	3639
QY	1741	CAGATCATCTCTGAGTCTGGCAGAGACCCAAAGAGAGGCCAAGATTTCAGAGACTT	1800
Db	3640	CAGATCATCTCTGAGTCTGGCAGAGACCCAAAGAGAGGCCAAGATTTCAGAGACTT	3699
QY	1801	GTGGAAAGTATAGTGATTTGCTTCAGCAATATGTACACAGAGTGTAAAGACTACATC	1860
Db	3700	GTGGAAAGTATAGTGATTTGCTTCAGCAATATGTACACAGAGTGTAAAGACTACATC	3759
QY	1861	CCAAATCAATGCCATCTACTGACAGGAATATAGTGGGTTTACATCTCAACTCTGCTCTCT	1920
Db	3760	CCAAATCAATGCCATCTACTGACAGGAATATAGTGGGTTTACATCTCAACTCTGCTCTCT	3819
QY	1921	GAGGACTTTCTCAAGGAAGTATTTACGCTCCGAGTTTAAITCAGGAAGCTCIGATGAT	1980
Db	3820	GAGGACTTTCTCAAGGAAGTATTTACGCTCCGAGTTTAAITCAGGAAGCTCIGATGAT	3879
QY	1981	GTCAATATGTAATAGCTTTCAAGTTTATGAGCTTGGAGAAATCAAAATCTTTGAGAA	2040
Db	3880	GTCAATATGTAATAGCTTTCAAGTTTATGAGCTTGGAGAAATCAAAATCTTTGAGAA	3939
QY	2041	CTTTTACGAGTGCACCTCTGATTTGATGATCTACAGGGCAGCAGATCTCTGTC	2100
Db	3940	CTTTTACGAGTGCACCTCTGATTTGATGATCTACAGGGCAGCAGATCTCTGTC	3999
QY	2101	G 2101	
Db	4000	G 4000	
RESULT 2			
LOCUS	AX481481	4017 bp	DNA
DEFINITION	Sequence 95 from Patent WO02055693.	linear	PAT 16-AUG-2002
ACCESSION	AX481481		
VERSION	AX481481.1	GI:22316395	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		

BASE COUNT 1236 a 915 c 903 g 963 t
ORIGIN

Query Match	99.8%	Score 2097.8	DB 9	Length 4017	
Best local Similarity	99.9%	Prod. No. 0			
Matches 2099	Conservative	0	Mismatches 2	Indels 0	Gaps 0
Qy 1	TAATACAGATGTGCCAAATGGGTTTCATGTTAACTTGGNAATAATGCCGCGGAGCA 60				
Db 1651	TAATACAGATGTGCCAAATGGGTTTCATGTTAACTTGGNAATAATGCCGCGGAGCA 1710				
Qy 51	GAGGACCTGAAACTGTCTTCACAGTTAAACAAGTCTTATACAGAGACGTTACTTGGATT 120				
Db 1711	GAGGACCTGAAACTGTCTTCACAGTTAAACAAGTCTTATACAGAGACGTTACTTGGATT 1770				
Qy 121	TTACTCGGACGAGTTAATACAGACAAATGCACTACAGTAATAGCAACGAAATAATGCGC 180				
Db 1771	TTACTCGGACGAGTTAATACAGACAAATGCACTACAGTAATAGCAACGAAATAATGCGC 1830				
Qy 181	ATCACTAAGGAGACGACGTCATCTCTTAACTTACCATCATGAATGTTTCCCTGCGAAGAT 240				
Db 1831	ATCACTAAGGAGACGACGTCATCTCTTAACTTACCATCATGAATGTTTCCCTGCGAAGAT 1890				
Qy 241	TCAGGACCTATGCTTCAGACGCGAGGAATGTATACACAGGGGAGAAATCCTCCAGAG 300				
Db 1891	TCAGGACCTATGCTTCAGACGCGAGGAATGTATACACAGGGGAGAAATCCTCCAGAG 1950				
Qy 301	AAAGAAATTAACAATCAGAGATCAGGAAGCACCATTCTCTCGAAACCTCAGTGATCAC 360				
Db 1951	AAAGAAATTAACAATCAGAGATCAGGAAGCACCATTCTCTCGAAACCTCAGTGATCAC 2010				
Qy 361	ACAGTGGCCATACAGAGTTCACGACCTTTAGACTGTCTAGTAAATGGTGTCCCGAGCCT 420				
Db 2011	ACAGTGGCCATACAGAGTTCACGACCTTTAGACTGTCTAGTAAATGGTGTCCCGAGCCT 2070				
Qy 421	CAGATCACTTGGTTTAAACAACACACAAAATACACAAGGCGCTGGAAATTAATTAGGA 480				
Db 2071	CAGATCACTTGGTTTAAACAACACACAAAATACACAAGGCGCTGGAAATTAATTAGGA 2130				
Qy 481	CCAGGAGCAGCAGCGTGTATTGAAGAGTCACAGAGAGGATCAAGGTGTCTATCAC 540				
Db 2131	CCAGGAGCAGCAGCGTGTATTGAAGAGTCACAGAGAGGATCAAGGTGTCTATCAC 2190				
Qy 541	TGCAAGCCACCAACACAGAGGCGCTCTGTGGAAAGTTACGATACCTCACTGTTCRAGA 600				
Db 2191	TGCAAGCCACCAACACAGAGGCGCTCTGTGGAAAGTTACGATACCTCACTGTTCRAGA 2250				
Qy 601	ACCTCGGACAAAGTCTAATCTGGAGCTGATCACTTAACATGACCTGTCTGGGTGGACT 660				
Db 2251	ACCTCGGACAAAGTCTAATCTGGAGCTGATCACTTAACATGACCTGTCTGGGTGGACT 2310				
Qy 661	CTCTCTGGCTCTTATACCTCTCTATCCGAAATGAAAGGTCCTCTCTGAAATA 720				
Db 2311	CTCTCTGGCTCTTATACCTCTCTTATCCGAAATGAAAGGTCCTCTCTGAAATA 2370				
Qy 721	AAGACTGACTACCTATCAATTAATAGGACCCAGATGAAGTTCCTTTGGATGAGCAGTGT 780				
Db 2371	AAGACTGACTACCTATCAATTAATAGGACCCAGATGAAGTTCCTTTGGATGAGCAGTGT 2430				
Qy 781	GAGCGGCTCCCTTATGATGCCAGCAATGGGAGTTTCCCGGAGAGACATTAACCTGGCG 840				
Db 2431	GAGCGGCTCCCTTATGATGCCAGCAATGGGAGTTTCCCGGAGAGACATTAACCTGGCG 2490				
Qy 841	AAATCACTTGGAGAGGCGCTTTTGGAAAGTGTCTCAGCATCAGCATTTGGCATTAAG 900				
Db 2491	AAATCACTTGGAGAGGCGCTTTTGGAAAGTGTCTCAGCATCAGCATTTGGCATTAAG 2550				
Qy 901	AAATCACTTGGAGAGGCGCTTTTGGAAAGTGTCTCAGCATCAGCATTTGGCATTAAG 960				

```

QY 2041 CTTTACCGAATCCACCTCCATGTTGATGACTACCAAGGGGAGCAGCAGCAGCTGTTG 2100
|||||
Db 3691 CTTTACCGAATCCACCTCCATGTTGATGACTACCAAGGGGAGCAGCAGCAGCTGTTG 3750
|||||
QY 2101 G 2101
Db 3751 G 3751

RESULT 4
RAFL11
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

SOURCE
Rattus norvegicus (strain Fisher F344) adult lung and testis cDNA
to mRNA, clone 4-47-4/3-8.

ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (sites)
Yamane,A., Seetharam,L., Yamaguchi,S., Gotoh,N., Takahashi,I.,
Neufeld,G. and Shibuya,M.
A new communication system between hepatocytes and sinusoidal
endothelial cells in liver through vascular endothelial growth
factor and Flt tyrosine kinase receptor family (Flt-1 and
KDR/Flk-1)
Oncogene 9 (3), 2683-2690 (1994)
94336223
2 (bases 1 to 4734)
Shibuya,M.
Unpublished
3 (bases 1 to 4734)
Shibuya,M.
Direct Submission
Submitted (06-FEB-1994) Masabumi Shibuya, Institute of Medical
Science, University of Tokyo, Department of Internal Medicine;
4-6-1 Shirokanedai, Minato-Ku, Tokyo 108, Japan (Tel:03-5449-5550,
Fax:03-5449-3425)

FEATURES
source
location/Qualifiers
1..4734
/organism="Rattus norvegicus"
/strain="Fisher F344"
/db_xref="taxon:10116"
/cclone="4-47-4/3-8"
/tissue_type="lung and testis"
/dev_stage="adult"
1..4734
/gene="flt-1"
237..4247
/gene="flt-1"
/feature="flt-1"
/feature="extracellular domain: AAL-758, transmembrane
domain: AA759-780, tyrosine kinase domain: AA813-1152,
kinase insert region: AA930-994."
/codon_start=1
/product="Flt-1 tyrosine kinase receptor"
/protein_id="BA05857.1"
/db_xref="GI:500379"
/translation="MVSCWDIVLPLCALLCGLITGCGSKLKGPELSLKGTHVPMQ
AGQTLFLKRGGAHWSLPTIVSDKLSVTRACGRNRFQFCITLILMACANH
GLXSRFLPKSKKKEKWSALIEVSDAGSFIEHSDIPKLVHMTGREGHIIICRV
TSNITVILKPEPDALTPOGRIANDSRGFIHANATYKEIGLLTCEATVNGHYQT
SYLTHQTNILDOISPPRFLRGQTLVLNCTVITDLNIRVMSWNPCKAKRA
SIQRIDQSNHNVFHSVLKINNVSRDKGLYCRVSGSSFFENISVHYEKGFI
SVKRWQOETIAKGRSHRLSMKRAFSPVWLKDPVPEKSNRSPVLSLI
KDVIAEDAGDYLILGKQSKLFNLIATILVYKPKIYKSSVSSLPVFLGSRQ
VLTCIVYGIPOPTIKLWHPCNHNHKKERNECFGSEESFILDSSNINRIFGIOR
MMVIEGINKVSLVADSRIPGYSIKAFNKGIVTDFVDPVNGPFWLSLEKI

```

```

PTEGDLKLSGVSVKFLYRDITWILLTAVNRTHMSISKOKMATDYDSTITLNLVYK
NWSLEDGIVACRARNIVYIGELLKTEVLROLEAPLILLONLSHREVSIGSTIIDC
QARVPAPQIINEKNHKKIQEPGIIIGPNSLTFTERTVEDEGVSRCAINQKGV
ESSAYLTVQSTSDKSNLEIITUTCVAAATFWLLTLFIRLKRSEVKTIDYLSII
MDPEVFLDEQERLPYDASKWEFARERLKLGRGAGKVQVQASAFGKIKSPICR
TVAVMKI KEGATASEYKALMTLKLTHIGHLNVNLLGACTKGGPLVWYVYCKY
GNLSNLIKSRDFFCLINKDAALHMEPKKLEPDLPDQKPLDSVSSSEFTSSGQFQ
EDKSVSDVEGDEYSEISKOPLMEDLIYSFOVARGNEFISRRKCIHRDLAARNIIL
SENNVVKICDFGLARDIVKNPDYVVRGDIRLPLKMMNPESIFDKVYSIKSDVWSYGL
LWEIFSLGGSPYGVQMBDEFCSLKAGEMRIPETATPEIYQIMLDCMHKDPKPPR
FAELVEKLGLDLOANVQDQKADYIPLNAILTRNSGTYSVTFSEDFKDPGDKFPH
SGSDSDRVYNFAKFMESLERIKTFEELSPNATSMFEDYHLDTSLSLPLKRFATIE
IKKASMKIDLRITTSKSEAGLSLDLPSPFCISSGCHIRPVQDEDDDELKESCCS
PPPDYNSVLYSSPPA"
BASE COUNT 1271 a 1263 c 1209 g 991 t
ORIGIN
Query Match 76.1%; Score 1599.2; DB 10; Length 4734;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 1787; Conservative 0; Mismatches 313; Indels 0; Gaps 0;
QY 1 TATATCAGAGTGGCAAAAGGTTTCAATGTTAACTTGGAAAATAATGCCAGGAAGGA 60
Db 1887 TACGTCACAGATGGCAACGGCTTTCACGTTTCCTTGGAAAAGATACCAACCAAGGA 1946
QY 61 GAGGACCTGAACCTGCTTGCACAGTTAACAGTTTCTATACAGAGAGCTTACTTGCAT 120
Db 1947 GAGGACCTGAACCTGCTTGCAGTAAATTCCTGTACAGAGACATACCTGTGATC 2006
QY 121 TATCTGGGACAGTAAATACAGACAAATGCATACAGTATTAGCAAGCAAAAATGCC 180
Db 2007 CTGCTACGACAGTAAACACAGGACCAATGCACATAGCATCAGTAAGCAAAAATGCC 2066
QY 181 ATCACTAAGGACACTCCATCACTCTTAACTTACCAATCAATGTTTCCCTGCAAGAT 240
Db 2067 ACCACTCAGGACTTCCATCACTCTGACCTTCTCAAGATGTTCTGTGGAAGAC 2126
QY 241 TCAGGACCTTGCCTTCAGAGGACAGGAATGTATACAGGGGAGAAATCTCCAGGAAG 300
Db 2127 TCGGACCTTATGCTTCAGAGGACAGGAATGTATACAGGGGAGGAAGATCTCTCGGAAG 2186
QY 301 AAGAATAATACATCAGAGAGCAGGAGCAACCACTCTCTCCGAACCTCAGTATACAC 360
Db 2187 ACAGAAAGTCTCGTTAGAGATTTGGAAGGCCACTCTGCTTCAAAACCTCAGTGACCAC 2246
QY 361 ACAGTGGCATCAGCAGTTCACACCTTTAGACTGTCTATGTTGTGTCGCCGAGCT 420
Db 2247 GAGGTGTCATCAGTGGCTCCAGGACCTTAGACTGTCAAGCTAGAGGTGCTCGGCT 2306
QY 421 CAGATCACTTGGTTTAAACCAACCAATACAGAGAGCCTGGAATTTATTAGGA 480
Db 2307 CAGATCACTTGGTTTCAAAACACCAATACAGAGAGCCTGGAATTTATTAGGA 2366
QY 481 CCAGGAGCAGCAGCTGTTTATTTGAAGAGTTCACAGAGAGGATGAAGTGTCTATCAC 540
Db 2367 CCAGGAAACACAGCCTGTTTATTTGAAGAGTTCACAGAGAGGATGAAGTGTCTATAGG 2426
QY 541 TCAAGAGCCACCAACCAAGAGGCTCTGTGGAAGTTCAGCATACCTCAGTCTTCAAGGA 600
Db 2427 TCCGAGGCCACCAACCAAGAGGCTCTGTGGAAGTTCAGCATACCTCAGTCTTCAAGGA 2486
QY 601 ACTCGGAAAGTCTTAATCTGGAGGTGATCACTCTACATGCCACCTGTGTGCTGCGACT 660
Db 2487 ACCTCAGACAGTCAAACTTGGAGTGTACCTCAGTGTGCTGCTGCTGCGACA 2546
QY 661 CTCCTCTGGCTTATTAACCTCTCTTATCGGAAAATGAAAAGGTCTTCTTCTCAATA 720
Db 2547 CTCCTTGGCTCTTCTTAACTCTCTTCTATCGGAAACTGAGCGGTCTCTTCTCCGAGTA 2606
QY 721 AAGACTGACTACCTATCAATATATGACCCAGATGAAGTTCCTTGGATGACAGTGT 780
Db 2607 AAGACGACTACCTGTCAATATCAATGACCCAGATGAAGTTCCTTGGATGACAGTGT 2666

```

QY 781 GAGCGGCTCCCTTATGATCCAGCAAGTGGAGTTTCCCGGAGAGACTTAACTGGSC 840
 DB 2657 GAAAGCGTGGCTATGATCCAGCAAGTGGAGTTTCCCGGAGAGACTTAACTGGSC 840
 QY 841 AATACACTTGGAGAGGGGCTTTTGGAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 2727 AATACACTTGGAGAGGGGCTTTTGGAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 901 AATACACTTGGAGAGGGGCTTTTGGAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 2787 AATACACTTGGAGAGGGGCTTTTGGAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 961 GAGTCAAAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 2847 GAGTCAAAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1021 GTGGTAACTTGGAGAGGGGCTTTTGGAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 2907 GTGGTAACTTGGAGAGGGGCTTTTGGAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 1081 TACTGCAAAATGAAATCTTCCCACTTACCTCAAGCAAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 2967 TACTGCAAAATGAAATCTTCCCACTTACCTCAAGCAAAAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1141 AACAGCAATGAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3027 AACAGCAATGAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1201 CAAGCAAAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3087 CAAGCAAAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1261 TTTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3147 TTTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1321 AAGGAGGCTTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3207 AAGGAGGCTTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1381 GAGTCTGCTTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3267 GAGTCTGCTTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1441 TCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3327 TCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 1501 CCGGATGATGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3387 CCGGATGATGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1561 ATCTTGGCAAAATCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3447 ATCTTGGCAAAATCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1621 GAAATCTCTCTTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3507 GAAATCTCTCTTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1681 AGTGGCTTGGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3567 AGTGGCTTGGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1741 CAGATCATGCTGAGTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3627 CAGATCATGCTGAGTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1801 GTGGAAGTCTGAGTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 DB 3687 GTGGAAGTCTGAGTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900
 QY 1861 CCAATCAATGCTGAGTCTGAGGAAGTCTGATGAGTGGTTCACAGCATCAGCAATTTGCAATAG 900

DB 3747 CCCTCAACCCCACTAGTACAGCAAGTGGTTCACAGCATCAGCAATTTGCAATAG 3806
 QY 1921 CAGGACTTCTCAGCAAGTATTTAGCTCCCAAGTATTTAATTCAGCAAGTCTGATGAT 1980
 DB 3807 GAGGACTTCTCAGCAAGTATTTAGCTCCCAAGTATTTAATTCAGCAAGTCTGATGAT 1980
 QY 1981 GTGAGTATGTAATGCTTTCAGTTCATGAGCTTGAAGAAATCAAAACCTTTGAGAA 2040
 DB 3867 GTGAGTATGTAATGCTTTCAGTTCATGAGCTTGAAGAAATCAAAACCTTTGAGAA 2040
 QY 2041 GTTTTACCAATGCTTTCAGTTCATGAGCTTGAAGAAATCAAAACCTTTGAGAA 3926
 DB 3927 GTTTTACCAATGCTTTCAGTTCATGAGCTTGAAGAAATCAAAACCTTTGAGAA 2100
 RESULT 5
 MUSRKB
 LOCUS
 DEFINITION Mus musculus receptor tyrosine kinase (FLT) mRNA, complete cds.
 ACCESSION L07297
 VERSION L07297.1
 KEYWORDS receptor protein tyrosine kinase.
 SOURCE Mus musculus (strain C57BL/6J, sub_species domesticus) cDNA to mRNA.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Finney, H., Kelleher, K., Morris, G.E., Bean, K., Merberg, D.M., Kriz, R., Morris, J.C., Sockdeo, H., Turner, K.J. and Wood, C.R.
 Molecular cloning of murine FLT and FLT4
 Oncogene 8 (8), 2293-2298 (1993)
 9330572
 8393164
 Location/Qualifiers
 1..6055
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /sub_species="domesticus"
 /db_xref="taxon:10090"
 1..6055
 /gene="FLT"
 /27..4028
 /gene="FLT"
 /note="putative"
 /codon_start=1
 /product="receptor tyrosine kinase"
 /protein_id="AAA40078.1"
 /db_xref="GI:293783"

AGTFLKRCGEAHSNSPTIVSCDRLSITPPSACRONROPCSTILLITDIAQNH
 TGLYCRYLPTSTSKKRAESSITFVSDAGSPFEMTDIPKLVHMEGRQILIPCR
 VSPNVTILKFPEDLIPDQRITWDSRRGFLIANATYKELGLCEATVNGHLYQ
 INVLHRONTILDVQIPSPVRLHGTOLVNGTITELNTRVQNSNPKATKR
 ISYKRRDASHNNVHSLKINNVESDKGLTCRVKSSQSFNTSVHVEKGR
 KAVITFDAGDYILLIKOSRFLKNTATILVNVKQIYKSVSSLPPLPIGSR
 QVLCTVYGPRTITLWVHCHHNSKERYDFCTENESFIDPSNLSNLSIESIQ
 EMVTEGNTKTVSLVADSDPTGYSCHAEKNGIGVPERNKIYFIDVPSHVSLEK
 KNSVEDSKLSCVNVKFLYRDTILKTLVTRVNRTHHSISKQKATIDYSLNLI
 CARGVAPOLINEKNWKTQPPGILIPGNSTILFVETVEDEGVRCATNOKGA
 VPSAATVOGTSDKSNELILITCTCAATLFWLLTLFIRKIPSSSEKTYDLSI
 IVDPEVLDEQERIPYDASKWEFERLKLGLGFGKGVQASAGFIRKSPIC
 RTAVKMLKGAASEKALMTLILHIGHLNVNLLGCTKOGGLMWIYVEYCK
 YNSLNYLKRDLFCNLDAALHMLKESLEGLQEQKFLDSVSSSVISSEFP
 EDRSVODEGDEYSEISKOPTMEDLTSYFQVARGMEFLSSRCHIRDLARNIL
 SENWVKICDFGLARDIKYNDYVRGDTIRLPLKMAPEIPIKYSITKSDKSYGL
 IWEIFSAGGSPYGVOMDEDCSLKCEMRMPEYATPEIYOIMLDCWKHKDKEPRP
 FAEIVKGLQLOANQQQGDYIPLNAILIRNSSFTYSTFTFSDFLKFQADPHFP
 SGSDDDVRYNAFKFMSLERIKITFEELSPNSTSMFEDYDITSLGSPULKRFWTE
 TKPASKIDLRISKSEKNSLDLPRFSCFSSCGHTRPVQDDESELGESCSPPP


```

QY 2041 CTTTACCGAATGCCACCTCCAGTTCAGTACATACAGGCGAGCAGCAGCACTCTGTC 2100
||||| ||||||| ||||||| ||||||| ||| ||| ||||| ||||||| |||
Db 3717 CTTTACCGAATGCCACCTCCAGTTCAGTACATACAGCAGCAGCACTCTGTC 3776

QY 2101 G 2101
|
Db 3777 G 3777

RESULT 6
D88689
LOCUS D88689 6275 bp mRNA linear ROD 30-JUL-2002
DEFINITION Mus musculus mRNA for flt-1, complete cds.
ACCESSION D88689
VERSION D88689.1 GI:2809068
KEYWORDS
SOURCE Mus musculus (strain:C57/B) lung cdna to mRNA, clone_l1b,mouse lung
cdna clone:MD-9-11.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Finnerty,H., Kelleher,K., Morris,G.E., Bean,K., Merberg,D.M.,
Kriz,R., Morris,J.C., Sookdeo,H., Turner,K.J. and Wood,C.R.
TITLE Molecular cloning of murine FLI and FLI4
MEDIINE Oncogene 8 (8), 2293-2298 (1993)
PUBMED 9330572
REFERENCE 2
AUTHORS Kondo,K., Hirtuka,S., Subbalakshmi,E., Matsushima,H. and
Shiuya,M.
TITLE Genomic organization of the flt-1 gene encoding for Vascular
Endothelial growth factor (VEGF) receptor-1 suggests an intimate
evolutionary relationship between the 7-Ig and the 5-Ig tyrosine
kinase receptor
JOURNAL Gene (1998) In press
REFERENCE 3 (bases 1 to 6275)
AUTHORS Kondo,K.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-1996) Kunio Kondo, The Institute of Medical
Science, The University of Tokyo, Department of Clinical Oncology:
4-6-1 Shirokanedai, Minato-Ku, Tokyo 106, Japan
(Tel:-81-3-5449-5632, Fax:+81-3-5449-5428)
FEATURES
source
1..6275
/organism="Mus musculus"
/strain="C57/B"
/db_xref="taxon:10090"
/map="561-G3"
/clone="MD-9-11"
/tissue_type="lung"
/clone_lib="mouse lung cdna"
1..6275
1..315
/number=1
/evidence=experimental
1..251
252..4253
/function="receptor for VEGF and PlGF"
/standard_name="fms-like tyrosine kinase 1"
/codon_start=1
/evidence=experimental
/product="flt-1"
/protein_id="BAA24498.1"
/db_xref="GI:2809068"
/translation="MYSQWDIAVLFALLGCLLIGVSGSKLVPEI.SLKGIOHVMQ
AQLILKAGEAHSLSPIIVSQEDKRLSIPPSAGDNRQFCS.LITDIAQAH
TGLTYRLTSKRKKAESSIYIFVSDAGSPFIEMHIDIPKLVM.EGROLIPCR
VTSRNVILTKPFDFLTLDQGRFIINDSRGFIANATYKEGLIN.EATVNGHLVQ
TNLTHPOINTILDVOIRPSPVRIHGOILVNCTALIELATRVOMSNVPGRAIR
ASIRQIDRSHSNVHSHVSLKINNVESRDKGLYCRVKSQSFQSTVSHVTEKGI
ISVHRKQPQVEITAGRRSIRLSMKVKAFFSPPEIWLKDGSPAILKSARYLHVGISLI

```

```

IKDVTIEDAGDYITLIGIKOSRLFKMLTATILVNVKPOIYKSVSLSPSPPLPLGSR
QVLTIVYGIPIRPTIITLWHPCHHNSKRDFTENESEFILDPSNLGNRIEISIQ
RMVTEGNTKVTSTLVVADSOQPGIYSCRAFNKIGTVERNKEVYTDVDPNGFHSLEK
MFAEGEDLKSCVNVKELYEDITWILLRTYANRTHHSISKQKMAITQDYSITLNDVI
KNVSLSDSTYACARNIYTCEDILKRTIVLRDSEAPHLLQNSDYEVSISGTTLD
CGARGVPAPQIIFKNHKTQDEPGIILPGNSTLFTIERYEEDGVEYCRAINKGA
VMSAAYLVQIGISDKSNLELITCTVAAATLFWLLTLFIKRLKSSSEVKIDYLSI
IMDPEVPLDQECERLPYDASKWEFARLKLGLRGAGFAGVQVQSAFAGIKKSPIC
RTVAVKMLKEGATASEYKALMTLKLILTHIGHILNVNLLGACTKOGCPWLVNFKYK
YGNLSNYLKSREDLFCINKDAALHMLAKESLEPLGQCGOKPRLDSSVSSSVSSPFP
EDRSYDVEGDEYSEISKOPTLMDLISYFQVARGMEFLSKKCIHRDIAARNILL
SENWVKICDEGLARDIKNPYVRKRGDTRPLKAWMAPESIFDKYISKSDSWYSGVL
LWEIFSLGSGPYPGVQWDEDFCSRKEGMRMTPEYATPEIYQIMLDCHWKDPKRP
FAELVEKLGDLQANVQQCKDYIPLNAILTRNSGFTYSTPESDLFKCGFADPHF
SGSSDDRVYVNAFKFMSLERIKTFEELSPNSTSMFEDYQJDTSLILGSPILKRFNTE
TKPKASMXIDLRIASKSKAGLSDLPSPFCFSSCGHIRVQDDESELKESCCSPFP
DYSNVLYSSPPA"
316..412
/number=2
/evidence=experimental
413..542
/number=3
/evidence=experimental
543..757
/number=4
/evidence=experimental
758..930
/number=5
/evidence=experimental
931..1057
/number=6
/evidence=experimental
1058..1242
/number=7
/evidence=experimental
1243..1360
/number=8
/evidence=experimental
1351..1530
/number=9
/evidence=experimental
1531..1590
/number=10
/evidence=experimental
1591..1805
/number=11
/evidence=experimental
1806..1914
/number=12
/evidence=experimental
1907
/citation=[1]
/replace="t"
1915..2223
/number=13
/evidence=experimental
2224..2370
/number=14
/evidence=experimental
2371..2502
/number=15
/evidence=experimental
2503..2609
/number=16
/evidence=experimental
2610..2742
/number=17
/evidence=experimental
2743..2847
/number=18
/evidence=experimental
2848..2961

```

```

/number=19
/evidence=experimental
2952..3050
/number=20
/evidence=experimental
3051..3204
/number=21
/evidence=experimental
3205..3302
/number=22
/evidence=experimental
3303..3425
/number=23
/evidence=experimental
3426..3537
/number=24
/evidence=experimental
3538..3637
/number=25
/evidence=experimental
3638..3743
/number=26
/evidence=experimental
3744..3886
/number=27
/evidence=experimental
3792
/citation=[1]
/replace="a"
3886
/citation=[1]
/replace="c"
3887..3971
/number=28
/evidence=experimental
3972..4066
/number=29
/evidence=experimental
4067..5275
/number=30
/evidence=experimental
4115
/citation=[1]
/replace="g"
4254..5275
4464
/citation=[1]
/replace="c"
4517..4520
/citation=[1]
/replace="ccc"
5156
/citation=[1]
/replace="a"
5492
/citation=[1]
/replace="aaacct"
5954
/citation=[1]
/replace="g"
5954
polya_site
BASE COUNT 1665 a 1622 c 1584 g 1404 t
ORIGIN

```

```

Query Match      74.9%  Score 1574.5;  DB 10;  Length 6275;
Best Local Similarity 84.7%  Pred. No. 0;
Matches 1779;  Conservative 0;  Mismatches 319;  Indels 3;  Gaps 1;

QY  1  TATATCACAGATGTCCTAAATGGTTTCACTTGAATAAATGCGGCGGANGA 60
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1905 TACGTACAGATGCGGATGGCTTTCACGTTTCCTGGAAGATGCCAGCCGANGA 1964
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  61  GAGGACCTGAACCTGCTTTCACAGTTTACAAAGTCTTATACAGAGAGCGTTACTTGGATT 120
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db  1965 GAGGACCTGAACCTGCTTGTGGTGAATAAATTCCTGTACAGAGACATTACCTGGATT 2024
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  121 TACTGGGACAGTTAATACAGAACTACCTACAGTATTAGCAAGCAAAATGGCC 180
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2025 CTGCTACGGACAGTTAACAACAGAACCATGCCCATAGTATCAGCAAGCAAAATGGCC 2084
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  181 ATCACTAAGGAGCAGTCCATCACTCTTAATCTTACCATCATGAATGTTTCCCTGCAAGAT 240
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2085 ACCACTCAAGATTACTCCATCACTCTGAACCTTGATCAAGAGCTGCTCTAGAGAC 2144
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  241 TCAGGCACCTATGCTGAGAGCCAGGAATGTATACAGGGGGAATAATCTCCAGAAG 300
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2145 TCGGCACTATGCTGAGAGCCAGGAACATATACAGAGGGAAGACATCTTCGGAAG 2204
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  301 AAGAATAATACATCAGAGATCAGGAAGCACCATACCTCTCTGCGAAACCTCAGTATAC 360
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2205 ACAGAAGTCTCGTTAGAGATTCGGAAGCCACACCTCTTCAAAACCTCAGTACTAC 2264
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  361 ACAGTGGCAATCAGCAGTTCCACCCTTTAGACTGTATGCTGTATGTTGCCCGAGCT 420
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2265 GAGTCTCCATCAGTGGCTCTACGACCTTAGACTGTCAAGCTAGAGGTGTCGCCGCT 2324
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  421 CAGATCAGTTGGTTTAAACCAACACACAAATACACAGAGCCTGGAATATTTAGGA 480
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2325 CAGATCAGTTGGTTCAAAACACACCAAAATACAAAGAACCCGGAATATTTAGGA 2384
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  481 CCAGGAGCAGCAGCTGTTTATGAAGAGTACAGAGAGGAGTGAAGTGTCTATCAC 540
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2385 CCAGGAAACAGCAGCCTGTTTATGAAGAGTACAGAGAGGAGTGAAGTGTCTATAGG 2444
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  541 TGCAGGACCCACCAACAGAGGCTCTGTGGAAGTTCAGCATACCTCAGTGTTCAGGA 600
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2445 TGCAGGACCCACCAACAGAGGCTGTTGGAAGGCGGCTACCTACCTCCGTCGAGGA 2504
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  501 ACTCGGACAGTCTAATCTGGAGCTGATCACTTAACATGCACTGTGTCGCTCGACT 660
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2505 ACTCAGACAAAGTCAAACTTGAGCTGATCAGCTCAGTGCACATGCGTGGCTGCGACC 2564
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  661 CTCTCTGGCTCCTATTACCTCTCTTATCGGAAATATGAAGTCTTCTTCTGMAATA 720
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2565 CTCTTTTGGCTCTCTAATCTCTTATCAGAAACTGAAGCGGTCTTCTTCCAGTA 2624
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  721 AAGACTGACTACCTATCAATTAATGAGCCACAGATGAATTCCTTTGGATGAGCAGTGT 780
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2625 AAGACAGACTACCTGATCAATTAATGAGCCACAGATGAATTCCTTTGGATGAGCAGTGT 2684
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  781 GAGCGCTCTCTTATGATCCAGCAAGTGGAGTTTCCCGGGGAGAGACTTAAACTGGCC 840
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2685 GAACGCTGCTCTATGATGCCAGCAAGTGGAGTTTGCAGGGAGAGACTGAAACTAGGC 2744
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  841 AAATCACTTGSNAGGGGCTTTTGGAAAGTGGTTCAGCATCAGCAATTTGGCAITTAAG 900
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2745 AAATGGCTCGGAAGAGGGGCTTTTGGAAAGTGGTTCAGCACTTGCATTTGGCATTAAG 2804
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  901 AAATCACTTACCTCCCGGACTGTGCTGTGAATAATGTGAAGAGGGGGCCACGGCCAGC 960
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2805 AAATCACTTACCTCCCGGACTGTGCTGTGAATAATGTGAAGAGGGGGCCACGGCCAGC 2864
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  961 GAGTACAAAGCTCTGATGACTGAGCTAAATATTTGACCCACATTTGGCCACCTCTGAAC 1020
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2865 GAGTACAAAGCTCTGATGACTGAGCTAAATATTTGACCCACATTTGGCCACCTCTGAAC 2924
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  1021 GTGTTAACTCTGGAGCCTTGCACCAACAGCAGGAGGCGCTCTGATGGTATGTTGAA 1080
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2925 GTGTTAACTCTGGAGCCTTGCACCAACAGCAGGAGGCGCTCTGATGGTATGTTGAA 2984
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  1081 TACTGCAATATGGAATCTCTCAACTACCTCAAGAGCAACAGTACTTATTTTCTC 1140
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  2985 TACTGCAATATGGAATCTCTCAACTACCTCAAGAGCAACAGTACTTATTTTCTC 3044
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  1141 AACAGAGTGCAGCAGTCTACATGAGCCTTAAGAGCAAGAAAAATGGAGCCAGCCTGGA 1200
    ||||||||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Mon Dec 9 10:33:37 2002

Db 2045 AACAGGACCGACCTTTCATATGAGCTCAGAAGACAGAGCGCTGACAGGCGCTGGAG 3104
 Qy 1201 CAAGGTAAAGACAGATTAAGCTACAGGAGTGAAGCTTTCCAGGCTCCGC 1260
 Db 3105 CAGGCGCAGAGCGCGCTGACAGTGT---CAGCAGCTCAAGTGTACAGGCTCAGC 3161
 Qy 1261 TTTCAGGAGATTAAGCTGAGTGTGTAGGAGACAGAGGATTTGACGCTTTTAC 1320
 Db 3162 TTCCCTGAGACCGAGCGTGTAGGATGTGAGGAGACAGGATTTACAGTGTAGATCTCC 3221
 Qy 1321 AAGGACCCATCATATGAGGATCTGATTTCTTACAGTTTCAAGTGTGCGAGGAGG 1380
 Db 3222 AAGCAGCCCTCAGCAGGAGAGCTGATTTCTTACAGTTTCAAGTGTGCGAGGAGG 3281
 Qy 1381 GAGTTCCTCTTCAGGAGGATGATTCATGCGGACCTTGTAGGAGGAGGAGGATTTT 1440
 Db 3282 GAGTTCCTCTTCAGGAGGATGATTCATGCGGACCTTGTAGGAGGAGGAGGATTTT 3341
 Qy 1441 TCTGAGAACAGCTGTGTAGGATTTGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1500
 Db 3342 TCTGAGAACAGCTGTGTAGGATTTGATTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1560
 Qy 1501 CCCGATATGTGAGAACAGCTGTGTAGGATTTGATTTGGCTTGGCTTGGCTTGGCTTGG 1560
 Db 3402 CCGTATATGTGAGAACAGCTGTGTAGGATTTGATTTGGCTTGGCTTGGCTTGGCTTGG 1620
 Qy 1561 ATCTTGACAAATCTACAGACCAAGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGT 1620
 Db 3462 ATCTTGACAAATCTACAGACCAAGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGT 1680
 Qy 1621 GAAATCTTCTTCTAGT 1680
 Db 3522 GAGATCTTCTTCTAGT 1740
 Qy 1681 AGTCCCTGAGGAGGAGT 1740
 Db 3582 AGCCGCTGAGGAGGAGT 1800
 Qy 1741 CAGATCATGT 1800
 Db 3642 CAAATCATGT 1860
 Qy 1801 GTGAAAGAGT 1860
 Db 3702 GTGAAAGAGT 1920
 Qy 1861 CCAATCAATGT 1920
 Db 3762 CCAATCAATGT 1980
 Qy 1921 GAGGAGT 1980
 Db 3822 GAGGAGT 2040
 Qy 1981 GTGAGATGT 2040
 Db 3882 GTGAGATGT 2100
 Qy 2041 CTTTACGAGT 2100
 Db 3942 CTTTACGAGT 2160
 Qy 2101 G 2101
 Db 4002 G 4002
 RESULT 7
 Locus M. musculus Empr2 cDNA.
 DEFINITION X78568
 ACCESSION X78568.1 GI:510654
 VERSION

Empr2 gene: receptor kinase.
 Mus musculus.
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 5482)
 Chai, K., Wall, C., Hanratty, R. and Keller, G.
 Isolation of a gene encoding a novel receptor tyrosine kinase from
 differentiated embryonic stem cells
 Oncogene 9 (4), 1261-1266 (1994)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1..5482
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_type="embryonic stem cells"
 249..4241
 /gene="Empr2"
 249..4241
 /gene="Empr2"
 /codon_start=1
 /product="embryonic receptor kinase"
 /protein_id="CAA5531.1"
 /db_xref="GI:5106558"
 /db_xref="MGI:195558"
 /translation="MVSQDIAVLFPYALLCGLLLGYGGGKLVPELSLKGTHQVMO
 AGQFLKCRGEAHSNLPITVQEDKRLSIPPSACGRONROFCTSLTLDTAQANH
 TGLTICRLPISKSKKKAESIVFSDAGSFIEMHTDIPKLVHTEGROLICRV
 TSPNVTILAKFPFDLIPDORITWDSRRGFIANAITYKEIGLLCNCVANGHLYOTN
 IYTHQNTIILDVIRPSPPLGALIGOTLVNCTATTELTVQWMSWYNGKATKAS
 IQRDRSHNNVHFHVLKINNVESRDKGIYTCRVKSGSQSPQSNVHYEKGFLI
 VKHRQVQETTAGRSYRLMKVAFSPETVWLKDGSPATLKARLYVHGLIILK
 DVTIEDAGDTILLGKQSLFKNLATILVNVKPOIYKESVSLPPLPLGSGRY
 LICVYGIPRETIWLWHPCHHNSKERYDECTEENESTLDPSSNLGNRIEYSQM
 TIVGINTKTVLWADSPGTPIYSCRAFNGKIGIVERNIKFYIYDVPNGFHVLEKMP
 ASGDKLSCVWVKLYRDTIWLITVNNRTHNSIKONGHSSYSITLILVKNV
 SLEDSGYACHARNIYTGEDILRKTELVRDSEAPHLLOLSDYEVSIGSTLIDCOA
 RGVAPQITIKNNKHIOQEPGILLGPGNSTLFIETVTEDEGVYRSTNOKGAVES
 AAVLVQGTSDKSNLELITITCVAAFLWLLTLFIKLRSSSEVKIYLSIIMD
 PVEVDEOCERLPYDASNEFARERKLKGLSLRGAGVQVQASAGFKIGSPICRV
 AVKMLKGAASEYKALMTELKILTHIGHLVNVLNGLACTKGGPLMVIYCYGN
 LSNYLSKRDJLCLNDAALHMLKKELEPGLGEOQKPRUDSVSSSVTSSFPEDR
 SVSDVEGDEDSKOPLMEDLISYFQVARMFESSKCIHRDLAARNILISE
 NVYKICDFGLARDIKNPDYVRGDTPLKWMAPESIFDKVYSTKSDVNSGVLWE
 IFSLGSPYPGVQDEDEFCRLKGNMRTPEVATPEIYQIMLDCWHDKPERFAE
 LVEKLGLQANVQDQKDIPLNAILIRNSFTYSTPTFSEDRGKGFADPHFSGS
 SDDVYVNAFKFMSLERIKTEELSPNSTSMFEDYOLDITSLIGSPLLKRFETETKP
 KASKWIDLRITASPRKAGLSLPRPFCFSCGHIRPVODDESELGKESCCSPEDYN
 SWLYSSPPA"
 BASE COUNT 1464 a 1446 c 1369 g 1203 t
 CDSIGIN
 Query Match 73.0%; Score 1534.2; DB 10; Length 5482;
 Best Local Similarity 84.1%; Pred. No. 0;
 Matches 1767; Conservative 0; Mismatches 328; Indels 6; Gaps 3;
 Qy 1 TATATCAGATGTGCAAAATGGGTTTCATGTAACTTGGAAAAATGCGACGAGGA 60
 Db 1896 TACGTACAGATGTGCGAAGTCTTTCACGTTCTCTTGGAAAAAGATCCAGCGGAGGA 1955
 Qy 61 GAGGACCTGAAACTGTCTTCACAGTTACAGAGTTCTTATACAGACCTTACTTGGATT 120
 Db 1956 GAGGACCTGAAACTGTCTTGGTTCATTAATTCCTGTACAGACATTAACCTGGATT 2015
 Qy 121 TTACTCGGACAGTATATACAGAACATCCACTACATATACAGCAAAAATGGCC 180

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE

1 Yamaguchi, S., Iwata, K. and Shibuya, M.
Soluble Flt-1 (soluble VEGFR-1), a potent natural antiangiogenic
molecule in mammals, is phylogenetically conserved in avians
Biochem. Biophys. Res. Commun. 291 (3), 554-559 (2002)
21845444
2 (bases 1 to 4272)
Shibuya, M.
Direct Submission
Submitted (11-JUL-2001) Masabumi Shibuya, University of Tokyo,
Institute of Medical Science, 4-6-1 Shirokane-dai, Minato-ku, Tokyo
108-8639, Japan (E-mail: shibuya@ims.u-tokyo.ac.jp)
Tel: 81-3-5449-5550, Fax: 81-3-5449-5425

FEATURES

source

1..4272

/organism="Gallus gallus"

/db_xref="taxon:9031"

/tissue="lib-whole embryo"

1..4272

/gene="vegfr-1"

257..4250

/gene="vegfr-1"

/note="flt-1

fms-like tyrosine kinase-1"

/codon_start=1

/product="vascular endothelial growth factor receptor-1"

/protein_id="BAB4690.1"

/db_xref="GI:18652861"

/translation="MPQLLSGLVILGAAFLIAGLSTGSKI KYPVLI SVK3PROHVVOAG
QTLNLTCEGMLHSLPEALSLSKRI NVI KYAGRNIGTFCFSHILSRJOANDIGR
YSCRPISPEVKKRESIVYVINDTSPNEVMSHSDIPKILIMTVGEMILKRGVTPY
IAVLTKIPREILIPDGKXILWDMRGFRLEAFYRIFGLISLSETIIGHKXSKTILI
HRETNIDIKLSLPRILKLGSLAINCIWKAANIRVMUWITPGEMKRGSVTQ
RDKTLEAVANFYSLIYIDKVRDIDKQYACHVKVSGPSNKI VNTIIVIVDKRFINLR
RKTMLAAGYKIRLPMKVAEPFVNLKGLPAAYFCARYVKNYSII IKDVA
EEDAGNIIILSLQWNLKSLJVLKVNKQIYENAVSFPNMLNLSLQKVLIC
TVGIPPKIITIMWPCRONHSLTRFGCSRTDGSFNKLTGSMNIGLQSLIERTAIL
EKNKTSATLVAAKSSGILYSCVASKVKAERNVSLFATDVPSFHLSEKVEIEG
ENLVSCANFHKDISWILPRIVQITKANKLKEYS ILIILIRVSLAHSGIY
ICHANISFTEVLORNDKSIQAEAPALLQIMPOTVNTSNAMIECOVHPIPEPI
TVFKNHEIQOESGIIIGPGRMLFIERYKEDEQLYOCIAINLKGVSFSTAYIVQ
TVERNSLEILITICVAAFLFWLLIIFIRKLKRPVSE KTNHLSIIMPDEVPL
DEQCECLPYDASKWEIARERLKGLSLGHGAGKVOASAFGLKSPICIVAVMLK
EGATAVEYKALMTELKILHGHSHINVLNIGACTKGGELMIVIECKGNLSYLK
SRNFSFKPDSLOGELMDKKGIEPVEGKKRLASVTSSESFSGFQEDKLSLDA
EEDDEDAELKPLIMEDLISYFQVARGMEFLSSRKCTFEDLAAIRLSENKVK
ICDFGLARDIYNPDYVRKGDARLPKMWAPESIFDKIYN KSDVRSYVLLWEIFSL
GASPYGVQIDEDFCSLKREGTRMRAPQATPEIYOIMLWTRSNPNERPSELVKR
LQDLQASVQGDGRDIPLDIIFIAESQFPASDPLCNKIPVPSFNCSTIERARYIN
TKIKPPORIKTEFELPKFKNVNDYQADSCWVLAPELFRLINTGSKYKTLFGWK
GVSRKSESLGSIKIPRSFCSCQLSESRRTYIGNTVLENKRAKCHSPDPYSSVV
HNSQPSI"

BASE COUNT 1266 a 1001 c 1083 g 922 t

ORIGIN

Query Match 55.6%; Score 1168.8; DB 5; Length: 4272;
Best Local Similarity 74.1%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 507; Indels 39; Gaps 5;
QY 3 TATCACAGATGCGCAATGGGTTTCATGTTAAGTGGAAATAATGCCGACCGAAGGAGA 62
Db 1907 TGTACACATGATACCAAGTGGATCCACATAGTTTGGNAAGAGTCCCATGGANGAGA 1966
QY 63 GCACCTGAACCTGCTGTGACAGAGTTTACAAAGTCTTATACAGAGAGCTTACATGATTT 122
Db 1967 GAATCTGGTATGCTGTGTCAGCCCAACAAATTCATGACAAAGACATTCCTGGATTTT 2026
QY 123 ACTCGGACAGTTAATACAGACAAATGCATACAGTAATAGCAACAAATAATGGTCAT 182
Db 2027 ACCGAGGACAGTACTATACCAACA-----AAAGCAAGGACCGCC 2068

QY 183 CACTAAGAGACATCCATCACTCTTAATCTTACATCATGAATGTTTCCCTSCAAGATTC 242
Db 2069 CAACAAGAGTACTCCATCACTCTCACTCTCACTCACCANTCAGGAATGTTTCCCTGGCACACT 2128
QY 243 AGGCACCTATGCTCTGAGAGCGAGGAATGTATACAGAGGGGAGAAATCCCTCCAGAGAA 302
Db 2129 AGGCACCTATACCTGCGGAGGAGGAACATATTCCGCGGAAGAAGTCTTTCAAAAGAA 2188
QY 303 AGAATATACATCAGAGATCAGAGACCACTACTCTCTCGGAACCTCAGTGATCACAC 362
Db 2189 AGACGTTTCAATCAGAGCGCAAGAGSCACCACTCTCTGCGGACGCTCATGGATCAGAC 2248
QY 363 AGTGGCCATCAGAGTTCACACCACTTTACAGTGTCTATGTAATGGTGTCCCGAGCCCTCA 422
Db 2249 GGTGAACACCACTTCTGTCATCTGCGAGTGCAGGTTTCATGCGATCCCGAGGCCCA 2308
QY 423 GAICACITGGTTTAAACAACACACAAATACAAAGAGCGCTGGAAATTTATTAGGACC 482
Db 2309 GATACCTGGTTTAAACAACAGAGGAATACAGCAAGATCAGGAATATTTTAGGGCC 2368
QY 483 AGGAAGCAGCAGCTGTTTATTGAAGAGTTCACAGAAGAGGATGAAGTGTCTATCACTG 542
Db 2369 TGGAGCGGAATGCTGTTTATTGAAGAGTAAAGAGGAGGAGGAGGAGGACTTTATCAGTG 2428
QY 543 CAAGCGCACCAACAGAGAGGCTCTGTGAAAGTTCAGCATACCTCAGTGTTCAGGAAC 602
Db 2429 CATGCGCACCACTTGAAGGGGTCTGTGGAGGACACAGCGTATGTCACAGTACAGGCAC 2488
QY 603 CTCGCGCAAGTCTAATCTGAGCTGATCACTTAACATGACCTGTGTGGTGGCGACTCT 662
Db 2489 AGTGGAGAGATCGAAGCTGAGCTGATAACCTGACCTGTACCTGTGTGGTGTGCAAGCT 2548
QY 663 CTTCGGCTCTTATTAACCTCTCTTATCCGAAAATGAAAAGGTCT---TCTTCGAAAT 719
Db 2549 TTCTGGCTCTGTCGACCTCTTCTATCGGAAGCTTACGAGGCGCTTACTTTCTGAAC 2608
QY 720 AAAGAC---TGACTACTTCAATTAATTAATGCCGCCAGATGAAGTTCCTTTGGATGACA 776
Db 2609 AAAGACGATCATACTCTGTAATCATCATGATCGGAGGAGTGCCTTTAGATGAGCA 2668
QY 777 GTGTGAGCGCTCCCTTATGATGCCAGCAAGTGGAGTTTGCCTGGGAGAGACTTAAACT 836
Db 2669 ATGTGAATGTCTGCTTACGACCGCAGCAAGTGGAGATTGACCGGAACGCTGAAGCT 2728
QY 837 GGCNAATCAGTTGGAGAGGGGCTTTTGGAAAGTGGTGTGAAGCATCAGCAATTGGCAT 896
Db 2729 AGGAAGTCACTTGGACATCGGCGCTTCCGGAAGGTGTACAGGCGCTCTGCTTTTGGAAI 2788
QY 897 TAAGAAATCACTACCTACGCGGACTGTGGCTGTCAAAATCTGAAAGAGGGGCCACGCG 956
Db 2789 TAAGAAATCGCAACATGCAAGATTTGTTGCTGTGAAAATCTGTAAGAGGGGCCACAGC 2848
QY 957 CAGCGAGTACAAGCTCTGATGACTGAGCTAAAATCTTGACCCACATTTGCCACCATCT 1016
Db 2849 GAGTGAGTACAAAGCACTGATGACTGAGCTCAAAATCTGATCCACATCGGCCATCATCT 2908
QY 1017 GAACGIGTTAACTGCTGGAGCGCTGCACCAACAGAGAGGGCGCTCTGATGGTGAATGT 1076
Db 2909 GAAATGTGNAATTTGCTGGAGCGCTGCACAAAATGATGGAGGCGCTCTGATGGTGAATGT 2968
QY 1077 TGAATCTGCAATATCGAAATCTCTCCAACCTACCTCAAGAGCAACAGTGAATTTT 1136
Db 2969 IGNAATCTGCAAGTACGGNAATTAACAATCTCTGAAAAGCAGAGGAGTCTTTTTCAG 3028
QY 1137 TCTCAAGAGTGCAGCACTACATGAGCCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1196
Db 3029 CCCCACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3088
QY 1197 GSACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1256
Db 3089 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3148
QY 1257 CGGCTTTCAGGAAGATAAAAGTCTGAGTGAT---GTTGAGGAAGAGGAGGAGTCTGACGG 1313


```

* 129384 129483: gap of 100 bp
* 129484 131747: contig of 2264 bp in length
* 131748 131847: gap of 100 bp
* 131848 132316: contig of 469 bp in length
* 132317 132416: gap of 100 bp
* 132417 134455: contig of 2039 bp in length
* 134456 134555: gap of 100 bp
* 134556 135527: contig of 972 bp in length
* 135528 135627: gap of 100 bp
* 135628 189051: contig of 53424 bp in length
* 189052 189151: gap of 100 bp
* 189152 189476: contig of 325 bp in length
* 189477 189576: gap of 100 bp
* 189577 191375: contig of 1799 bp in length
* 191376 191475: gap of 100 bp
* 191476 201473: contig of 9998 bp in length
* 201474 201573: gap of 100 bp
* 201574 202307: contig of 734 bp in length
* 202308 202407: gap of 100 bp
* 202408 204978: contig of 2471 bp in length
* 204979 204978: gap of 100 bp
* 204979 213331: contig of 8553 bp in length
* 213332 213631: gap of 100 bp
* 213632 218109: contig of 4478 bp in length
* 218110 218209: gap of 100 bp
* 218210 219800: contig of 1591 bp in length
* 219801 219900: gap of 100 bp
* 219901 220965: contig of 1065 bp in length.
FEATURES
    source
    Location/Qualifiers
        1. 220965
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="3"
        /map="3p21.3"
        /clone="RP6-32g23"
BASE COUNT 62577 a 44674 c 45661 g 65453 t 2600 others
ORIGIN

Query Match      45.5%; Score 956.4; DB 2: Length 220965;
Best Local Similarity 94.6%; Pred. No. 6.4e-268;
Matches 990; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1056 AGGCGCTCTGATGCTGATTTGAAATGCAAAATGGAATCTCTCCAACTACCTCAA 1115
Db 25255 AGGCGCTCTGATGCTGATTTGAAATGCAAAATGGAATCTCTCCAACTACCTCAA 25196

QY 1116 GAGCAACGTCACCTATTTTCTCAACAGGATGACAGCTACACATGGAGCGCTAGAA 1175
25195 GAGCAACGTCACCTATTTTCTCAACAGGATGACAGCTACACATGGAGCGCTAGAA 25136

QY 1176 AGAAAAATGGCCAGGCTCGAAGCAAGCAAGCAACCAAGCACTAGATAGCTCACCAG 1235
Db 25135 AGAAAAATGGCCAGGCTCGAAGCAAGCAAGCAACCAAGCACTAGATAGCTCACCAG 25075

QY 1236 CAGCGAAGTTTTCGAGCTCGGGTTTCAGAGAGATAAAAGTCTGAGTGTGAGGA 1295
Db 25075 CAGCGAAGTTTTCGAGCTCGGGTTTCAGAGAGATAAAAGTCTGAGTGTGAGGA 25016

QY 1296 AGAGCAGGATCTCAGCGTTTCTACAGAGCCCATCACTATGAGATCTGATCTTCTT 1355
Db 25015 AGAGCAGGATCTCAGCGTTTCTACAGAGCCCATCACTATGAGATCTGATCTTCTT 24956

QY 1356 CAGTTTTCAGATGCGCAGAGGATGAGTTCTGCTCTCTCCAGAAAGTCAITCATCGGGA 1415
Db 24955 CAGTTTTCAGATGCGCAGAGGATGAGTTCTGCTCTCTCCAGAAAGTCAITCATCGGGA 24896

QY 1416 CTGCGCAGGAGAAACATCTTTTCTCTGAGAAACAGCTGGTGAAGATTTGATTTGG 1475
Db 24895 CTGCGCAGGAGAAACATCTTTTCTCTGAGAAACAGCTGGTGAAGATTTGATTTGG 24836

QY 1476 CTTTCGCGCGGATATTTATAGAACCCCGATTTATGTCAGAAAGGAGATACCTGACCTCC 1535
Db 24835 CTTTCGCGCGGATATTTATAGAACCCCGATTTATGTCAGAAAGGAGATACCTGACCTCC 24776

```

```

QY 1536 TCTGAAATGGATGCTCCCGAATCTATCTTTGACAAAATCTACAGCAACCAAGACGACGT 1595
Db 24775 TCTGAAATGGATGCTCCCGAATCTATCTTTGACAAAATCTACAGCAACCAAGACGATGT 24716

QY 1596 GTGGTCTTACGGAGATATGCTGGGAAATCTCTCTTAGTGGGTCTCCATACCAGG 1655
Db 24715 GTGGTCTTACGGAGATATGCTGGGAAATCTCTCTTAGTGGGTCTCCATACCAGG 24655

QY 1656 AGTACAAATGGATGAGGACTTTTTCAGTCCGCTGAGGAAAGGATGAGGATGAGAGCTCC 1715
Db 24655 AGTACAAATGGATGAGGACTTTTTCAGTCCGCTGAGGAAAGGATGAGGATGAGAGCTCC 24596

QY 1716 TGAGTACTCTACTCTTGAATCTATCAGATCATCTGGACTGCTGGCACAGACCCAAA 1775
Db 24595 TGAGTACTCTACTCTTGAATCTATCAGATCATCTGGACTGCTGGCACAGACCCAAA 24536

QY 1776 AGAAAGGCCAAGATTTGCAAGACTTGTGAAACTAGTGGAAATCTGCTTCAGCAATGT 1835
Db 24535 AGAAAGGCCAAGATTTGCAAGACTTGTGAAACTAGTGGAAATCTGCTTCAGCAATGT 24475

QY 1836 ACAACAGATGTTAAAGACTTACATCCCAATCAATGCCATCTACAGCAAAATAGTGGTT 1895
Db 24475 ACAACAGATGTTAAAGACTTACATCCCAATCAATGCCATCTACAGCAAAATAGTGGTT 24416

QY 1895 TACATATCACTACTCTGCTTCTCTGAGAGCTTCTTCAAGAAAGTATTCAGCTCCGAA 1955
Db 24415 TACATATCACTACTCTGCTTCTCTGAGAGCTTCTTCAAGAAAGTATTCAGCTCCGAA 24355

QY 1956 GTTTAAATCAGGAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2015
Db 24355 GTTTAAATCAGGAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 24296

QY 2016 GGAAGAAATCAAAACCTTTTGAAGAACTTTTACCAGATTTTACCAGATTTTACCAGATTTT 2075
Db 24295 GGAAGAAATCAAAACCTTTTGAAGAACTTTTACCAGATTTTACCAGATTTTACCAGATTTT 24236

QY 2075 CCAGGGGAGCAGGAGCAGCTCTGTTGG 2101
Db 24235 CCAGGGGAGCAGGAGCAGCTCTGTTGG 24210

LOCUS AC104439 197279 bp DNA linear PRI 20-JUN-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-793E15, complete sequence.
ACCESSION AC104439 AC024739
VERSION AC104439.2 GI:21490240
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 197279)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Santhimachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197279)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 197279)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Santhimachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Jun 20, 2002 this sequence version replaced gi:1748621.
----- Genome Center

```

Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgchgs@u.washington.edu
 Drafting Center: UWGSC

----- Project Information

Center Project name: chr-3
 Center Clone name: RP11-793E15 (bc0564)

----- Summary Statistics

Sequencing vector: unknown; 52% of reads
 Chemistry: plasmid; 108752; 48% of reads
 Chemistry: Dye-terminator ET; 94% of reads
 Chemistry: Dye-terminator Big Dye; 6% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 197188 bases at least Q40
 Consensus quality: 197255 bases at least Q30
 Consensus quality: 197275 bases at least Q20
 Insert size: 197279; sum-of-contigs
 Quality coverage: 8.2x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP11-91E8 (UWGC:bc0216) AC026349
 3': CTD-2563A18 (UWGC:bc0730)

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

HindIII			BglII			EcoRI		
SeqDerMap	FrgrPrint	SeqDerMap	FrgrPrint	SeqDerMap	FrgrPrint	SeqDerMap	FrgrPrint	SeqDerMap
2687	2617	8949	8586	8596	8661			
6382	6410	2067	2160	6	<800			
512	<800	7846	7940	2742	2803			
449	<800	3734	3895	5376	5324			
2602	2763	1334	1301	1433	1478			
2590	2617	2287	2309	823	835			
8313	8291	1814	1918	1962	2002			

1711	1683	691	<800	2900	2954
9821	9472	5477	5348	1484	1478
516	<800	305	<800	1005	995
8587	8291	25245	25541	1181	1171
7446	7581	3988	4121	18560	19002
2088	2075	1633	1598	3603	3579
2509	2617	631	<800	4943	5076
3519	3501	90	<800	3239	3241
26	<800	402	<800	953	995
925	933	3350	3490	1621	1615
98	<800	4577	4515	6827	6900
6409	6410	1229	1183	79	<800
1383	1376	4221	4331	16418	16263
15048	15829	3980	4121	3361	3241
1876	1877	2548	2497	872	881
4169	4068	674	<800	4220	4193
1674	1683	2332	2309	2848	2803
52	<800	11445	11045	2283	2318
5227	5082	5692	5632	3211	3241
15464	15829	1385	1414	4615	4632
14333	14296	608	<800	6823	6900
1026	1026	3930	3895	886	881
545	<800	1465	1414	2032	2002
5621	6410	2382	2497	8680	8661
8413	8291	747	<800	5163	5076
1401	1376	1951	2012	1547	1478
1512	1471	642	<800	7058	6900
5801	5844	21060	21003	1647	1615
3256	3278	3700	3666	2753	2803
3448	3501	1321	1301	162	<800
2738	2763	10705	10502	1078	1093
237	<800	3128	3317	2156	2173
5715	5823	183	<800	5077	5076
201	<800	3255	3490	154	<800
1002	1026	5693	5632	1571	1478
221	<800	3380	3317	219	<800
2854	2899	237	<800	520	<800


```
----- 7317 431 ----- 1540 1478
----- 1145 2152 ----- 1478
----- 8291 249 ----- 1615
----- 1911 2599 ----- 1478
----- 3899 3847 695 ----- 9383
----- 5672 5644 41 ----- <800
----- 1711 1683 388 ----- 13115
----- 852 854 2500 ----- 4632
----- 2368 2389 2889 ----- 5076
----- 178 ----- 4430
----- 6589 ----- 6900
----- 3836 ----- 2803
-----

Query Match 45.4%; Score 953.2; DB 9; Length 197279;
Best Local Similarity 94.5%; Pred. No. 5.4e-267;
Matches 988; Conservative 0; Mismatches 58; Indels 0; Gaps 0:

QY 1056 AGGGCTCTGATGTTGTTGAATCTGCAATATGGAATCTCTCCAACTACCTCAA 1115
DB 68708 AGGGCTCTGATGTTGTTGAATATGCAATATGGAATCTCTCCAACTACCTCAA 68649
QY 1116 GACCAACGTGACATATTTTCTCAACAGGATGACGACTACATGAGCCCTAAGAA 1175
DB 68648 GAGCAATATGACATATTTTCTCAACAGGATGACGACTACATGAGCCCTAAGAA 68589
QY 1176 AGAAAAATGAGCGCGCTGGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1235
DB 68588 AGAAAAATGAGCGCGCTGGAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 68529
QY 1236 CAGCGAAGCTTTTCGAGCTCCGGTTTCAGGAAGATAAAGTCTGAGTGTGTTAGGA 1295
DB 68528 CAGCGAGAGCTTTTCGAGCTCCAAAGTTTCAGGAAGATAAAGTCTGAGTGTGTTAGGA 68469
QY 1296 AGAGGAGATCTCACCGTTTCTACAGAGCCCATCACTATGAGATCTCAATCTTIA 1355
DB 68468 AGAGGAGATCTCACCGTTTCTACAGAGCCCATCACTATGAGATCTCAATCTTIA 68409
QY 1356 CAGTTTTCAGTGGCCAGAGCAAGTTCCTGTTTCCAGAAAGTGCATTCATCGGGA 1415
DB 68408 CAGTTTTCAGTGGCCAGAGCAAGTTCCTGTTTCCAGAAAGTGCATTCATCGGGA 68349
QY 1416 CTGGCAGGAGACATCTTTTATCTGAGAACAGTGGTGAAGATTTGTAATTTGG 1475
DB 68348 CTGGCAGGAGACATCTTTTATCTGAGAACAGTGGTGAAGATTTGTAATTTGG 68289
QY 1476 CTTTCCCGGGATATTTTAAAGAACCCCGATTTATGAGAAAGAGGATACCTGCATCC 1535
DB 68288 CTTTCCCGGGATATTTTAAAGAACCCCGATTTATGAGAAAGAGGATACCTGCATCC 68229
QY 1536 TCTGAATGGATGGCTCCCGAATCTATCTTTGACAAAATCTACAGCAAGAGGAGCT 1595
DB 68228 TCTGAATGGATGGCTCCCGAATCTATCTTTGACAAAATCTACAGCAAGAGGAGCT 68169
QY 1596 GTGGTCTACGAGATTTCTGTTGGAAATCTCTCTTAGTGGGTCTCCATCCGAGG 1655
DB 68168 GTGGTCTACGAGATTTCTGTTGGAAATCTCTCTTAGTGGGTCTCCATCCGAGG 68109
QY 1656 AGTACAAATGGATGAGGACTTTTTCAGTGGCTGAGGAGGCAATGAGGATGAGGCTCC 1715
DB 68108 AGTACAAATGGATGAGGACTTTTTCAGTGGCTGAGGAGGCAATGAGGATGAGGCTCC 68049
```

```
QY 1716 TGAGTACTTACTCTCTGAAATCTATCAGATCATCTGCTGAGTGTGGCAGAGACCCAAA 1775
DB 68048 TGAGTACTTACTCTCTGAAATCTATCAGATCATCTGCTGAGTGTGGCAGAGACCCAAA 67989
QY 1776 AGAAGGCCAGATTTGCGAAGTTTGGAAAACCTAGGTGATTTGCTTCAAGCAAAATGT 1835
DB 67988 AGAAGGCCAGATTTGCGAAGTTTGGAAAACCTAGGTGATTTGCTTCAAGCAAAATGT 67929
QY 1836 ACAACAGGATGTTAAAGACTACATCCCAATCAATCCCATCTGACAGGAATAGTGGGT 1895
DB 67928 ATACAGAGATGTTAAAGACTACATCCCAATCAATCCCATCTGACAGGAATAGTGGGT 67869
QY 1896 TACATACTCAACTCTCTGCTTCTCTGAGGACTTCTTCAAGGAAAGTATTTTCAAGTCCGAA 1955
DB 67868 TACATACTCAACTCTCTGCTTCTCTGAGGACTTCTTCAAGGAAAGTATTTTCAAGTCCGAA 67809
QY 1955 GTTAAATCNSAGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2015
DB 67808 GTTAAATCNSAGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 67749
QY 2016 GGAAGAATCAAAACCTTTTGAAGAATTTTACCAGTATGCCACCTCCATGTTTCATCACTA 2075
DB 67748 GGAAGAATCAAAACCTTTTGAAGAATTTTACCAGTATGCCACCTCCATGTTTCATCACTA 67689
QY 2076 CCAGGGCCAGCAGCAGCAGCTCTGCTGG 2101
DB 67688 CCAGGGCCAGCAGCAGCAGCTCTGCTGG 67663

RESULT 11
AX481480
LOCUS AX481480 4044 bp DNA linear PAT 15-AUG-2002
DEFINITION Sequence 94 from Patent WO02055693.
ACCESSION AX481480
VERSION AX481480.1 GI:22316394
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kreutzer,R., Limmer,S., Rost,S. and Hadwiger,P.
TITLE Method for inhibiting the expression of a target gene
JOURNAL Patent: WO 02055693-A 94 18-JUL-2002;
Ribopharma AG (DE)
FEATURES
Source Location/Qualifiers
1..4044
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1163 a 885 c 1021 g 975 t
ORIGIN

Query Match 34.0%; Score 713.8; DB 6; Length 4044;
Best Local Similarity 65.5%; Pred. No. 3.8e-197;
Matches 1077; Conservative 0; Mismatches 562; Indels 5; Gaps 2:

QY 218 TCATGAATGTTTCTCCATCCAGATTCAGGCACTATGCTGCGAGCCAGGAATGATACA 277
DB 1886 TTAGATGCTATCTTTCAGGACCAAGAGACTATGCTGCTTCTCAAGACAGGAAGA 1945
QY 278 CAGGGAAGAATCTCTCCAGAGAAATTTACAATCAGAGATCAGGAAGCAACCATACC 337
DB 1946 CCAAGAAAGACATTCGCTGGTCCAGCAGCTCACAGTCTTAGACCGGTGGCACCACCA 2005
QY 338 TCTCGGAAACTCAGTATATCACAGTGGCCATCAGAGTCCACACTTTAGACTGTC 397
DB 2006 TCACAGGAACCTGAGATTCAGAGCAAGTATGGGAAAGCAATCGAAGTCTCATGCA 2065
QY 398 ATGCTAATCTGTCCTCCGAGCCTCAGATCAGTGGTTTAAACACACCAACAAATACAC 457
DB 2066 CGGCATCTGGGAATCCCTCCACAGATCATGTTGTTTAAAGATAATGAGACCTTGTAG 2125
```

QY 458 AAGAGCCTGGATTAATTAATAGACAGGAGCAGCAGCTGTTTATTAAGAGGTACAG 517
DB 2126 AAGATCAGCAATTAATAGAGTGGAAACCGGAACTTACATATCCGAGAGTGA 2185
QY 518 AAGAGATGAAGTGTCTATATCAAAAGGCACTAACACAGAGGCTCTGTGGAAGT 577
DB 2186 AAGAGAGGAGGAGGCTTACACTCCAGGAGTGTCTGTGCTGTGCAAGTGG 2245
QY 578 CAGCAATCTTCTATCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCT 637
DB 2246 AGGCAATTTTCTAATAGAGGAGGCTCTGAGCACTCTGAGCACTCTGAGCACT 2305
QY 638 CATGACCTGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 697
DB 2306 TAGGACAGGAGGCTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCT 2365
QY 698 TGAAGAGTCTCTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 754
DB 2366 TTAGCGGGGCAATGAGGAGGAGTGTGAGCACTCTGAGCACTCTGAGCACTCT 2425
755 ATGAAGTCTCTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCT 2485
QY 815 TTGCGCGGAGAGTCTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 874
DB 2486 TCCCGAGAGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2545
QY 875 TTAAGATCAGCAATTTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 934
DB 2546 TTAAGATCAGCAATTTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2605
QY 935 TGTGAAAGAGGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 994
DB 2606 TGTGAAAGAGGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2665
QY 995 TGACCCACATTTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1054
DB 2666 TCAATCAATTTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2725
QY 1055 GAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1114
DB 2726 GAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2785
QY 1115 AGAGCAAGTGTCTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1174
DB 2786 GAGCAAGAGAGTGTCTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2845
1175 AAGCAAAATGAGGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1234
2846 AAGATAGTGTGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2902
QY 1235 GAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1294
DB 2903 GTAGCAGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 2962
QY 1295 AAGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1354
DB 2963 AAGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 3022
QY 1355 ACAGTGTGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1414
DB 3023 ACAGTGTGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 3082
QY 1415 ACTGAGGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1474
DB 3083 ACTGAGGAGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 3142
QY 1475 GCTTGGCGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1534
DB 3143 GCTTGGCGGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 3202
QY 1535 CTCTAAATGAGTGTGAGTGTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACTCTGAGCACT 1594

RESULT 12
AR142803
LOCUS AR142803 4071 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6204011.
ACCESSION AR142803
VERSION AR142803.1 GI:15104089
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
1 (bases 1 to 4071)
Kendall, R.L., Mac, X., Thomas, K.A. and Tebben, A.
Human receptor tyrosine kinase, XDR
Patent: US 6204011-A 1 20-MAR-2001;
FEATURES
Location/Qualifiers
1..4071
BASE COUNT 1159 a 894 c 1025 g 982 t
ORIGIN

Query Match 34.0%; Score 713.8; DB 6; Length 4071;
Best Local Similarity 55.5%; Pred. No. 3.8e-197;
Matches 1077; Conservative 0; Mismatches 562; Indels 6; Gaps 2;

QY 218 TCAATGAATGTTTCCCTGCAAGTTCAGGCACTATGCTGCTGAGGAGGAGTGTATACA 277
DB 1886 TTAGAATGCACTCTTCCAGGAGGAGTGTGCTGCTGAGGAGGAGTGTATACA 1945
QY 278 CAGGAGGAGGAGTCTCCAGGAGGAGTGTATACAATCAGGATCAGGAGGAGGAGT 337
DB 1946 CCAAGAAAGAGCAATGCTGCTGAGGAGGAGTGTATACAATCAGGAGGAGGAGT 2005
QY 338 TCTGCGAACTCAGTGTATCAGACAGTGTGCTGAGGAGGAGTGTATACAATCAGGAGT 397
DB 2006 TCAAGAGGAGGAGTGTGAGTGTATCAGGAGGAGTGTGCTGCTGAGGAGGAGT 2065
QY 398 ATGCTAATGTTTCCCGAGGAGTGTATACAATCAGGATCAGGAGGAGGAGTGTATACA 457
DB 2065 CGGCACTGTGGAATTCCTGCTGAGGAGGAGTGTATACAATCAGGAGGAGGAGT 2125
QY 458 AAGAGCTGTGGAATTTTAGGAGGAGGAGGAGTGTATACAATCAGGAGGAGGAGTGTATACA 517
DB 2126 AAGAGCTGTGGAATTTTAGGAGGAGGAGGAGTGTATACAATCAGGAGGAGGAGT 2185
QY 518 AAGAGGAGGAGGAGTGTATACAATCAGGAGGAGGAGTGTATACAATCAGGAGGAGGAGT 577
DB 2186 AAGAGGAGGAGGAGTGTATACAATCAGGAGGAGGAGTGTATACAATCAGGAGGAGGAGT 2245
QY 578 CAGCAATCTGAGTGTGAGGAGGAGTGTATACAATCAGGAGGAGGAGTGTATACAATCTGAGT 637

Db 2246 AGCAATTTTCATTAATGAGGTCGCCAGGAAGACCACTTGGAAATCATATTTAG 2305
QY 638 CATGCACCTGTGTGCGACTCTCTTCTGGCTCTTATTAACTCTCTTATCCGAAAA 697
Db 2306 TAGGCACGGCGGTGATGCCAATGCTTCTGGCTATCTTCTGCTATCTTCTGCTAG 2365
QY 698 TGAAGAGTCT---TCTCTTGAATTAAGACTGATCTATCAATTAATTAAGACCCAG 754
Db 2366 TTAAGGGGCCAATGGAGGGGAACITGAAGACAGGTAGCTCTTCAATGCTATGATCCAG 2425
QY 755 ATGAAGTCTCTTGTGATGAGCAGTGTGAGCGCTCCCTTATGATGCGCAGCAAGTGGAGT 814
Db 2426 ATGACTCCATTTGATGACATTTGTAAGTACIGCTTATGATGCGCGTAATGGAGT 2485
QY 815 TTGCCCGGAGAGACTTAATGCTGGCAAAACACITGGAGAGGGCTTTTGGAAAAAGTGG 874
Db 2486 TCCCCAGAGACCGGTGAAGCTAGTAGAGATGAGCTCTGCTGGCGGTGGTCTTGGCAAGTA 2545
QY 875 TTCAAGCATCAGCAATTTGGCAATTAAGAAATACCTACCTACCTGGGAGCTGTGGCTGTGAAA 934
Db 2546 TTGAACAGATGCCCTTTGGAAATGACAAAGCAGCAACTTGCAGGACAGTAGCAGTCAAAA 2605
QY 935 TGTCTGAAGAGGGGCCACGGCCAGGAGTACAAAGCTCTGATGACTGAGCTAANAATCT 994
Db 2606 TGTGAAGAGAGGACCAACACAGAGTGAAGATGAGCTCTGATGATGACATCAGATCC 2665
QY 995 TGACCCACATTTGGCCACCATCTGAAGTGTATTACCTGCTGGAGCTGTGACCAAGCAAG 1054
Db 2666 TCAATCATATTTGGTACCATCTCAATGCTGCAAGCTTCAAGTGGCTGTGACCAAGCCAG 2725
QY 1055 GAGGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1114
Db 2726 GAGGGCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2785
QY 1115 AGGCAAAAGCTGACTTAATTTTCTCAAGAGGATGACAGTACACATGAGGACCTAAGA 1174
Db 2786 GGAGCAAGAGAAATGAATTTCTCCCTACAGAACCAAGGGGACGATTCGGTCAAGGA 2845
QY 1175 AAGAAAAATGAGCCAGGCTGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1234
Db 2846 AAGACTAGCTTGGAGCAATCCCTG---TGATCTGAAACGGCGCTTGGACAGCATCAACA 2902
QY 1235 GCAGCAAAAGCTTTGCGAGCTCCGGTTTTCAGGAAGATTAAGTCTGAGTGAATGAGG 1294
Db 2903 GTAGCCAGAGCTACAGCAGCTCTGATTTCTGGAGAGAGTCCCTCAGTGTAGAG 2962
QY 1295 AAGAGAGATTTCTAGCGTTTCTACAGAGCCCATCACTATGGAAGATCTGATTTCT 1354
Db 2963 AAGAGAGATCTCTGAAGATCTGATAAGTCTTCTGACCTTGGAGCATCTACTGT 3022
QY 1355 ACAGTTTTCAGTGGCCAGAGGATGAGTCCCTGCTTCCAGAAAGTGCATTCATCGG 1414
Db 3023 ACAGTTCAGTGGCTNAGGCAATGAGTCTTGGCAATCGGAAGTATCCAGAGG 3082
QY 1415 ACTGCGACGAGAAATCTTTATCTGAGAACAGTGGTGAAGATTTGATTTG 1474
Db 3083 ACCTGCGGCAAGAAATCTCTTATCTGAGAACAGTGGTGAAGATTTGATTTG 3142
QY 1475 GCCTTCCCGGGAATTTATAGAACCCCATTTATGAGAACAGGAGATCTGATCTTC 1534
Db 3143 GCTTGGCCCGGGAATTTATAGAACCCCATTTATGAGAACAGGAGATCTGATCTTC 3202
QY 1535 CTCTGAATGAGTGGCTCCGCAATCTATCTTGAACAAATCTACAGCAATCAAGCAGG 1594
Db 3203 CTTTGAATGAGTGGCCCGGCAAGAAATTTTGAACAGATGATCAATCAAGAGTGA 3262
QY 1595 TGTGCTTACGAGTATTTGCTGGGAATCTTCTCTTATGAGTGGTGTCTCCATACCCAG 1654
Db 3263 TCTGCTTTTGGTGTGTCTGCTGGGAATTTTCTCTTATGAGTGGTGTCTCCATACCC 3322
QY 1655 GAGTCAAAATGATGAGGACTTTTGGAGTGGCTGAGGAGAGGCAATGAGATGAGAGCT 1714

Db 3323 GGTAAGAATGATGAAGAATTTTGTAGGCGATTTGAAAGAGAACTAGATGAGGGCC 3382
QY 1715 CTGAGTACTACTCTCTGAAATCTATCATGATCATGCTGAGCTGTGCGACAGACCCAA 1774
Db 3383 CTGATATATATACACCAAGAAATGTACCAAGACCATGCTGAGCTGTGCGACGGGAGCC 3442
QY 1775 AGAAAGGCCAAGATTTGCGAAGCTTTGGAAGAACTAGGTGATTTGCTTCAAGCAATG 1834
Db 3443 CTCAGAGACCCAGCTTTTCAAGTGTGCTGAACATTTGGGAATCTCTTGAAGTAAATG 3502
QY 1835 TACAGCAGATGTAAGACTACAT 1859
Db 3503 CTCAGCAGGATGGCAAGACTACAT 3527
PFSLT 13
AP201382
LOCUS AR201382 4071 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6359115.
ACCESSION AR201382
VERSION AR201382.1 GI:20252270
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4071)
AUTHORS Rendall, R.L., Mao, X., Thomas, K.A. and Tebben, A.
TITLE Human receptor tyrosine kinase, XDR
JOURNAL Patent: US 6359115-A 19-MAR-2002;
FEATURES Location/Qualifiers
Source 1..4071
BASE COUNT 1169 a 894 c 1026 g 982 t
CPIGIN
Query Match 34.0%; Score 713.8; DB 6; Length 4071;
Best Local Similarity 65.5%; Pred No. 3.8e-197;
Matches 1077; Conservative 0; Mismatches 562; Indels 6; Gaps 2;
QY 218 TCAATGATTTTCCCTGCAAGATTCAGGACCTATGCTGCTGAGGCGAGGAAATGATACA 277
Db 1886 ITAAGATTCATCTTGCAGGACCAAGGAGACTATGCTGCTTGTCTAAGACGAGA 1945
QY 278 CAGGGAAGAAATCTCCAGAGAAAGAAATTAATCAATCAGAGATCAGGAGCAGCATACC 337
Db 1846 CCAAGAAAGACATTCGGTGGTCAAGCAGCTCACAGTCTTAGAGCGTGTGGCACCAG 2005
QY 338 TCTGGAAGACCTCAGTATGATCAGACAGTGGCCATCAGCAGTTCACACCTTTAGACTGC 397
Db 2786 ICACAGGAACCTCGAGATCAGGACACANGTATTTGGGAAAGCTGANGTCTATGCA 2085
QY 398 ATGCTAATGCTGCCCCAGGCTCAGATCACTTGGTTTAAAGAACCAACCAATACAC 457
Db 2785 CGCATCTCGGAATCCCTCCACAGATCAATGTTTAAAGATTAATGAGACCTTTAG 2125
QY 458 AAGAGCTCGGAATATTTAGGACCAAGGACGACGCTGTTTATGAAAGAGTCAAG 517
Db 2126 AAGACTCAGCATGATTAATGAGGATGGGACCGGACCTCACTATCCGACAGTGA 2185
QY 518 AAGAGATGAAGTGTCTATCACTTCGAAAGCCCAACCAAGAGGGCTCTGTGGAAGTT 577
Db 2196 AAGAGCAAGAGGCTCTACACTGCGCAGGATGAGTGTCTTGGCTGTGCAAGATG 2245
QY 578 CAGCATACCTCAGCTTTCAGGAGAACCTCGGACAGTCTAATCTGGAGCTGATCACTTAA 637
Db 2245 AGCATTTTTCATTAATAGAGGTGCCAGGAAAGCAGGACTTGGAAATCATTTATCTAG 2305
QY 638 CATGACCTGTGGCTGCGACCTCTCTTCTGGCTCTAATTAACCTCTCTTATCCGAAA 697
Db 2305 TAGGCACGGCGTGTATGCCATGTTCTTCTGCTGCTCTTCTGCTATCTCTACGACCG 2365
QY 698 TGAAGAGTCT---TCTCTTGAATTAAGACTGATCTATCAATTAATTAAGACCCAG 754

Db 2366 TTAAGCGGGCAATGGAGGGAACTGAACAGAGGTAACCTGCTCCATGCTGATGATCCAG 2425
 QY 755 ATGAAGTCTCTTTGGATGAGCAGTGTGAGCGGTCCCTTATGATGCGCAAGTGGAGT 814
 Db 2426 ATGAACCTCCCATTTGGATGAACAATTTGAAAGAGTCTGCTTATGATGCGCAAAATGCGAAT 2485
 QY 815 TTGCGGGGAGAGACTTAAATCTGGCAAAATCACTTGGAGAGGCGCTTTTGGAAAGTGG 874
 Db 2486 TCCACAGAGACCGCTGAAGCTAGTGTAGCTCTTTGGCGCTGGTCTTGGCCAACTGA 2545
 QY 875 TTCAAGCATCAGCATTTGGCAATTAAGAAATCACTAGTGGCGGAGTGTGGCTGTGGAA 934
 Db 2546 TTGAAGCAGATGCTTTTGAATTTGACAGACACACACTTGCAGGACAGTAGCAGTCAAA 2605
 QY 935 TCGTGAAGAGGGCGCCAGCCAGCAGTACAACTCTGATGACTGAGTCTAAAATCT 994
 Db 2606 TGTGAAGAGAGGCAACACAGTGAAGTGAAGTCTCACTGCTCACTCAAGATCC 2665
 QY 995 TGACCCACATTTGGCCACCACTGAACCTGTAACTGTAACTGTGGAGCTTGCACACAG 1054
 Db 2666 TCATTTCATTTGCTCACCATCTCAATGTGGTCAACCTTCTAGGTGCTGTACCAAGCCAG 2725
 QY 1055 GAGGCGCTCTGATGTGATTTGTAATCTGCAAAATATGAAATCTTCCAACTACTCA 1114
 Db 2726 GAGGCGCTCTGATGTGATTTGTAATCTGCAAAATATGAAATCTTCCAACTACTCA 1174
 QY 1115 AGACAAAGCTGACTTATTTTCTCAACAGATGCGAGCTACACATGAGCCTTAAGA 1174
 Db 2786 GGACAGAGAGAAATGAATTTGCTCCCTACAAAGCAACCAAGGCGCAGTCCGTCAAGGGA 2845
 QY 1175 AGAAGAAATGAGAGCGCTGGAACAGGCAAGCAACCAAGAGACTAGATAGGCTACCA 1234
 Db 2846 AAGACTACCTTTGGAGCAATCCCTG---TGGATCTGAACGCGCTTGGAGCACTACCA 2902
 QY 1235 GCAGCGAAAGCTTTGCGAGCTCGGCTTTACAGAGATAAAAGTGTGATGTGAGG 1294
 Db 2903 GTACCCAGAGCTACAGCAGCTCTGGAATTTGGAGGAGAGTCTCCAGTGTGAGG 2962
 QY 1295 AAGAGAGAGTCTTCAAGGTTTCTCAAGAGCCCATCACTATGGAAGATCTGATTTCT 1354
 Db 2963 AAGAGAGAGTCTTCAAGGTTTCTCAAGAGCCCATCACTATGGAAGATCTGATTTCT 1354
 QY 1355 ACAGTTTCAAGTGGCCAGAGGATGAGTTCCTGCTTCCAGAAAGTGCATTAAGGG 1414
 Db 3023 ACAGTCTCAAGTGGCCAGAGGATGAGTTCCTGCTTCCAGAAAGTGCATTAAGGG 1414
 QY 1415 ACCTGGCAGGCAACATCTTTTATCTGAGACACAGTGTGAGATTTGTGATTTG 1474
 Db 3083 ACCTGGCAGGCAACATCTTTTATCTGAGACACAGTGTGAGATTTGTGATTTG 1474
 QY 1475 GCTTGGCCGGATATTTATAAGAACCCCGATTATGTGAGAAAGAGATACCTGACCT 1534
 Db 3143 GCTTGGCCGGATATTTATAAGAACCCCGATTATGTGAGAAAGAGATACCTGACCT 3202
 QY 1535 CTCTGAATAGTGTCTCCGAACTATCTTTGACAAATCTACAGCAACCAAGAGAGG 1594
 Db 3203 CTCTGAATAGTGTCTCCGAACTATCTTTGACAAATCTACAGCAACCAAGAGAGG 3262
 QY 1595 TGTGGCTTACAGGATTTCTGTGGGAAATCTTCTCTAGTGGTCTCCATACCCAG 1654
 Db 3263 TGTGGCTTACAGGATTTCTGTGGGAAATCTTCTCTAGTGGTCTCCATACCCAG 1654
 QY 1655 GAGTACAAATGGAAGAGACTTTTTCAGTCTGAGGGAAGCAATGAGGATGAGGCTC 1714
 Db 3323 GGGTAAAGATTTGAAGAAATTTGTAGGCGATTGAAGAGAGACTAGATGAGGCGCC 3382
 QY 1715 CTGAGTACTCTCTCTGAACTATCAGATCACTGAGTGTGCTGAGCAGAGAGCCAA 1774
 Db 3383 CTGATTTATACACAGAAATGATACAGACCACTGCTGAGTGTGCTGAGCAGAGAGCCAA 3442
 QY 1775 AAGAAAGGCAAGATTTGCAAGATTTGGAAGAACTAGTCTATTTGCTTCAAGCAATG 1834
 Db 3443 GTCAGAGACCCAGTTTTCAGAGTTGGTGAACATTTGGGAAATCTCTTGAAGATG 3502

QY 1935 TACACAGGATGCTAAGACTACAT 1859
 Db 3503 CTCACAGGATGGCAAGACTACAT 3527
 RESULT 14
 AX481483
 LOCUS
 DEFINITION Sequence 4071 bp DNA linear PAT 16-AUG-2002
 ACCESSION AX481483
 VERSION AX481483.1 GI:22316397
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.
 TITLE Method for inhibiting the expression of a target gene
 JOURNAL Patent: WO 02055693-A 97 18-JUL-2002;
 Ribopharma AG (DE)
 FEATURES
 Location/Qualifiers
 source 1. 4071
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1159 a 894 c 1025 g 983 t
 ORIGIN
 Query Match 34.0% Score 713.8; DB 6; Length 4071;
 Best Local Similarity 65.5% Pred. No. 3.8e-197;
 Matches 1077; Conservative 0; Mismatches 562; Indels 5; Gaps 2;
 QY 218 TCATGATGTTTCCTCAAGATTTCAGGCACCTATGCTGCGAGCCAGCATGTATACA 277
 Db 1886 TTAAGAAATGCACTCTTCAGGACCAAGAGAGTATGCTGCTTGCCTCAAGACAGGA 1945
 QY 278 CAGGGGAAGAAATCCCTCCAGAAAGAAATTAATCAATCAGATCAGGAAGCAGCATACC 337
 Db 1945 CCAAGAAAGACATTTGGTGGTCAGGCAGCTCAGAGCTCTAGAGCGTGTGGCACCACGA 2005
 QY 338 TCTCGGAACCTCAGTATCAGATGCGCATCAGCTGCGCATCAGCTTCCACCATTGACTGC 397
 Db 2006 TCACAGGAACCTGGAGATCAGACGCAAGATTTGGGAAAGCATCGAAGTCTATGCA 2065
 QY 398 ATGCTAATGTTGCTCCGAGCTCAGATCCTTTGGTTTAAACAAACCAACAAATACAAC 457
 Db 2056 CGGCACTGGGAATCCCTCCACAGATCAITGGTTTAAAGATATGAGACCTTGTAG 2125
 QY 458 AAGAGCTGTGAATATTTTAGGACCGAGGACGACGCTGTTTATGAAAGATCAGAG 517
 Db 2126 AAGACTCAGGCAITGTTGAAGGATGGAAACCGAACCTCACTATCCGAGATGAGGA 2185
 QY 518 AAGAGATGAGGTGCTTACATGCAAGCCCAACCAAGAGGAGGCTCTGTGGAAGTT 577
 Db 2186 AAGAGGAGGAGGAGGCTCTACACTGCCAGGATGCACTGTCTTGGCTGTGCAAGTTG 2245
 QY 578 CAGCATACCTCAGTCTCAGGAACCTCGGACAGTCTTAATCTGAGCTGATCAGCTTAA 537
 Db 2246 AGGCAATTTTCAATATAGAAAGTGGCCAGGAAAGAGAACTTGGAAATCAITTTAG 2305
 QY 538 CATCAGCTGTGGCTGCGACTCTCTTCTGGCTGCTATTAACCTCTTATCCGAAATA 697
 Db 2306 TAGCAGCGGGTGATGCGCAITGTTCTTCTGGCTACTTCTTGTCAITCACTTACGAGCG 2365
 QY 628 TGAAGAGTCT---TCTTCTGAAATGAAGACTACTTACTATCAATATATATGAGCCAG 754
 Db 2366 TTAAGCGGCCAATGAGGAGGAGTGAAGACAGGCTACTTGTCCATGCTGATGATCCAG 2425
 QY 755 ATGAAGTCTCTTTGGATGAGCAGTGTGAGCGGCTCCCTTATGATCCAGCAAGTGGAGT 814
 Db 2426 ATGACATCCCATTTGATGAACATTTGGAACGAGTGTCTTATGATGCGAGCAATGGAT 2485

QY 815 TTGCCCCGAGAGACATTAACCTGGCAAAATCACTTGGAGAGCGGCTTTTGGAAAGTGG 874
 DB 2486 TCCCCAGAGCCGGCTGAGCTAGGTAGCTTCTTGGCCGTGGTCTTTTGGAAAGTGG 2545
 QY 875 TTCAGCATCAGATTTGGCATATAGAAATCACTTGGCCGAGCTTGTGCTTGGAAA 934
 DB 2546 TTGAAGCAGATGCTTTGGAATTTGACAGAGCAAGCACTTCCAGGACAGTAGAGCA 2605
 QY 935 TGCTGAAAGAGGGGCGCAGCGGAGTACAAAGCTCTGATGACTGAGCTTAAATCT 994
 DB 2606 TGTGAGAGAGAGGAGCAACACAGCTGAGCACTGAGCTCTTGAATCTTGAATCT 2665
 QY 995 TGACCCAGATTTGGCCACCATCTGACAGCTGTTTAACTTGGTGGCACTTCCAGCAAGCAAG 1054
 DB 2666 TCATTCATATTTGTCACCACTCAATGTTGTCACCTTCTAGTACCTTGGTACAGCAG 2725
 QY 1055 GAGGCGCTCTGATGCTGATTTGATGATGATGATGATGATGATGATGATGATGAT 1114
 DB 2726 GAGGCGCACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2785
 QY 1115 AGAGCAAAAGCTGATTTTCTCAACAGAGTACAGCACTTACAGTGGATCTAAGA 1174
 DB 2786 GGAGCAAGAGAAATGATTTGTCCTTACAGAGCAAGAGGCGCTTATCTGCTCAAGGA 2845
 QY 1175 AGAGAAATGAGAGCCAGCCCTGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1234
 DB 2846 AAGACTAGTTGAGCAATCCCTG---TGAATCTGAAGAGGCTGAGCAAGCAAGCAAG 2902
 QY 1235 GAGCAGAAAGCTTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1294
 DB 2903 GTAGCCAGCTTACGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2962
 QY 1295 AAGAGAGAGATTTGAGCGTTTCTACAGAGCCCACTACATGATGAGCAAGCAAGCAAG 1354
 DB 2963 AAGAGAGAGCTTCTGAGATCTGATGAGCACTTCTGAGCTTGGAGCACTTCTGCTG 3022
 QY 1355 ACAGTTTCAAGTGGCCAGAGCATGAGTTCCTTCTCCAGAAAGTGCATTTATCGGG 1414
 DB 3023 ACAGTTTCAAGTGGCTAGAGCATGAGTTCCTTCTCCAGAAAGTGCATTTATCGGG 3082
 QY 1415 ACCTGGCAGCAGAAATCTTCTTATCTGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 1474
 DB 3083 ACCTGGCAGCAGAAATCTTCTTATCTGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 3142
 QY 1475 GCTTGGCCGGGATTTTATAGAAACCCGATTTATGAGAAAGAGATACCTTGCCTC 1534
 DB 3143 GCTTGGCCGGGATTTTATAGAAACCCGATTTATGAGAAAGAGATACCTTGCCTC 3202
 QY 1535 CTCTGAATGAGTGGCTCCCGAATCTATCTTGAAGAAATCTAGCAGCAAGCAAGCAAG 1594
 DB 3203 CTCTGAATGAGTGGCTCCCGAAGAAATTTTGAAGAAATCTAGCAGCAAGCAAGCAAG 3262
 QY 1595 TGTGGCTTACGAGATTTGCTGTTGGAAATCTTCTGTTAGTGGTCTTCTCCATCCOAG 1654
 DB 3263 TGTGGCTTACGAGATTTGCTGTTGGAAATTTTCTGTTAGTGGTCTTCTCCATCCOAG 3322
 QY 1655 GAGTAAATGAGTGGATTTTGAAGTGGCTGAGAGAGGATGAGGATGAGGATCT 1714
 DB 3323 GGTAAAGATGAGGAAATTTTGAAGTGGCTGAGAGAGGATGAGGATGAGGATCT 3382
 QY 1715 CTGAGTACTTCT 1774
 DB 3383 CTGATATCT 3442
 QY 1775 AAGAAAGGCAAGATTTGCAAGCTTGTGGAAGAACTAGTGTGATTTGCTTCTCAAGCAAG 1834
 DB 3443 GTCAGAGCCGACGTTTTCAGGTTGTGGAGCAATTTGGAAATCTCTCTCAAGCAAG 3502
 QY 1835 TACAGAGATGTTAAAGACTACAT 1859
 DB 3503 CTCAGCAGATGCAAGACTACAT 3527

RESULTS
 AF063658
 LOCUS
 DEFINITION
 Homo sapiens vascular endothelial growth factor receptor 2 (KDR)
 mRNA, complete cds.
 AF063658
 VERSION
 AF063658.1 GI:3132832
 SOURCE
 Homo sapiens.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 4071)
 Y.Y., Whitney, R.G. and Sato, J.D.
 Coding region for human VEGF receptor KDR (VEGFR-2)
 JOURNAL
 Unpublished

REFERENCE
 2 (bases 1 to 4071)
 Y.Y., Whitney, R.G. and Sato, J.D.
 Direct Submission
 Submitted (05-MAY-1998) Adirondack Biomedical Research Institute,
 10 Old Barn Rd., Lake Placid, NY 12946, USA
 JOURNAL
 Location/Qualifiers

FEATURES
 source
 1..4071
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_type="endothelium"
 /issue_type="umbilical vein"
 1..4071
 /gene="KDR"
 1..4071
 /gene="KDR"
 /note="VEGFR-2; tyrosine kinase"
 /codon_start=1
 /product="vascular endothelial growth factor receptor 2"
 /protein_id="AAC16450.1"
 /db_xref="GI:3132832"
 /translation="MESKVLIAVALWCVETPRASVCLPSVDLPRIKIDILTIK
 ANILQITCGQRDLNEMNTNQSSEORVEVTECDGLFCKLTIPVIGNDIGAYK
 CYRIDEASVYVVDYDFEFIAVDSDHGVYITENKNTVWICPLSGISNLVS
 LCARPEKRFDPGNRISWDSKGTIPISYISYAGVFCARINDESGSIMYIVV
 VYRIYDVLSPSGIELSVGKLVNLARTIELNVDIDFNWEPSPKHOKLVNED
 KTGSGEMKFIETITIDGVTSDQGLYTCASSGLMTKNSITVRVHEKPFVAGS
 GMSVLAIVCEVRIPAKVICYPPEIKYKAGIPLESNTIKAGHVLITMEVSEHD
 TGNIVILINPISKQSHVSLVYVPPQIGKSLISPVDSYQYGTITUTITVAT
 PPHIHMYWLEECANEFSAQSVNTNPYCEEMRSVEDFGGKIEVKNQFALIE
 GKNITVTLVIAQANYSALYKCAVKNVGERVISTHTVTRPEITLOPDQPTQES
 VSLMCTADRTTFENLTKYLPQPLIHVGLPTPKCNLDLWLKNAIMFNSINDI
 LIMELKASLDQDQDYVCLAODRTKKRHCVRQLTVLERVAPITIGLNQITSGE
 SIENSCITASCNPPQIMWFKDNETLVEDSGTVLKDGNRLTIRVRKEDSLTQPC
 SVLSCAVEAFFIIEGAQEKINLEIILVGTAVIAMFFWLLVILIRIVKANGGEIK
 TGYLSVMDDELPHDECHERLPDASKWEPFRLKLGKPLGRGAGQVTEADAFGI
 DKATIRIVAVKMKEGATSEHRLMSELKILIHGHNLVNVNLLGACTKPGGLMV
 IVETCFGNLSYLRKNEFVYKTKGARKGQKDYGAIPDLKRLDSDITSSOSS
 ASSGFVEKSLSDVEEAPDLKYDLTLEHLICYFOVAKGMEELASRKLIRDLA
 ARNILLSEKNVWICDFGLARDYKDPDYVKGDARLPLKWAPEITFDRTVLOSIV
 WSPGVLLWEIFSLGASPYPCVAIDEECRRLKGTMRAPDITTEMOTMLDCWGE
 BSQPTSELSVNEHLNLLQANQQQDKDYIVLPISETLSMEEDSLSLPTSPSCME
 EYVQCPAFHNTAGISQTLQNSRKRSPSVKTFEDIPLEPEYKVPIDNQTDSGM
 VLASEELKIDREHRIKLSPSGFMVPSKSVASRNSQITSGTSGVSHSDIDITVVS
 SEAEELKLIEIGVTGTAOHPDSCTTILSPPV"

BASE COUNT 1169 a 894 c 1025 g 983 t
 ORIGIN

Query Match 34.0%; Score 713.8; DB 9; Length 4071;
 Best Local Similarity 55.5%; Pred. No. 3.8e-187;
 Matches 1977; Conservative 0; Mismatches 562; Indels 6; Gaps 2;

CY 218 TCATGAATGCTTTCCTGCAAGATTCAGGACCACTATCCCTTCAGAGCCAGCAATGTATACA 277
 DB 1886 TTAGATGCTTTCCTGCAAGCAAGGAGCATGTCTGCTGCTTCTCAAGCAGGAAGA 1945
 QY 278 CAGGGGAAGAAATCTCTCCAGAGAAAGAAATTAATACAGATCAGAGACCAATACC 337

Db	1946	CMAGAAAGACATTTGGTGTGTCAGGCAGCTCACAGCTCTAGAGCGTGTGGCACCCAGCA	2055
QY	338	TCCTCGAAACCTCAGTGTATCACAGAGTGGCCATCACAGCTTCCACACTTTTAGACTGTC	397
Db	2006	TCACAGAAACCTCGACAATTCAGACCACAGTATTTGGGAAACGATCGAAGTCTCATGCA	2065
QY	398	ATGCTAATGGTGTGCGCGGACCTCAGATCATTTGGTTTAAAAACAACCAACAAATACAAC	457
Db	2066	CGGCAATCGGGAATCCCTCCACAGATCATGTGGTTTAAAGATAATGAGACCTGTGTAG	2125
QY	458	AAGACCCCTGGAATTAATTTAGGACAGGAAACACAGACGCTGTATTAAGAGATCAAG	517
Db	2126	AAGACTCAGGCATTGTATTAAGAGTAGGGACCGGACCTCACTATCCGCGAGTAGGA	2185
QY	518	AAGACGATGAAGTGTCTATCTACTGCAAGACCAACACAGAGGCTCTGTGGAAAGTT	577
Db	2186	AGGAGGACAGGCGCTCTACGCTGCCAGCTCATCGAGTGTCTGGCTGCGAAAGTGG	2245
QY	578	CAGCACTCACTGTCTCAAGGACCTCGACAGCTCTAATCTGAGCTGTATCACTCTAA	637
Db	2246	AGGCCATTTTCATTAATAGAGGTGCCAGGAAAGAGGAACITGGAAATCATTAATCTAG	2305
QY	638	CATGCACCTGTGGCTCGGACTCTCTCTGCTCTCTATTAACCCCTCCCTTAACCGAAAA	697
Db	2306	TAGGCACGCGGGATGGCATGTCTCTGGTACTCTCTGTCATCTCTACCGACCG	2365
QY	698	TGAAAAGTCTCTCTCTGGAATTAAGACTCACTACCTATCAATTAATATGACCCAG	754
Db	2366	TTAAGCGGCGCAATGGACGGGAACTGAAGACAGGCTACTGTCCATGCTCATGTATCCAG	2425
QY	755	ATGAAGTTCCTTTGGATGAGCAGTGTGAGTGGCTCCCTTATGATGCCACGAGTGGAGT	814
Db	2426	ATGAACCTCCCATTTGGAATGAACATTTGTGAACGACTGCTTATCATGTACAGCAATCCAT	2485

Search completed: December 8, 2002, 11:59:20
Job time : 6045.36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 04:25:08 ; Search time 111.05 Seconds
(without alignments)
2048.198 Million cell updates/sec

Title: US-09-778-900a-5_COPY_500_600

Perfect score: 101

Sequence: 1 statggtgctgtgtgtctcc.....tatgtaaatgtttcaagtt 101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

tal number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SID2/gcgdata/geneseq/emb1/NA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/emb1/NA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/emb1/NA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/emb1/NA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/emb1/NA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/emb1/NA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/emb1/NA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/emb1/NA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/emb1/NA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/emb1/NA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/emb1/NA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/emb1/NA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/emb1/NA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/emb1/NA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/emb1/NA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/emb1/NA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/emb1/NA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/emb1/NA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/emb1/NA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/emb1/NA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/emb1/NA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/emb1/NA2001A.DAT:*
- 23: /SID2/gcgdata/geneseq/emb1/NA2001B.DAT:*
- 24: /SID2/gcgdata/geneseq/emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	101	100.0	1256	22	AAD19456	Human fit-1 gene D
2	32	31.7	288	23	AAS51865	Staphylococcus aur
3	32	31.7	300	19	AAV45103	Staphylococcus aur
4	32	31.7	300	19	AAV45098	Staphylococcus aur
5	32	31.7	300	19	AAV33271	Staphylococcus aur
6	32	31.7	300	20	AAV56688	S. aureus RAIC cod
7	32	31.7	303	23	AAS4756	Staphylococcus aur
8	32	31.7	303	23	AAS55130	Staphylococcus aur
9	32	31.7	303	23	AAS55306	Staphylococcus aur

C 10	32	31.7	3400	21	AA251225	Staphylococcus aur
C 11	32	31.7	13542	18	AAV74465	Staphylococcus aur
C 12	31.6	31.3	160771	24	AB088179	Human osteoblast d
C 13	31.4	31.1	36778	23	ABL07548	Drosophila melanog
C 14	30.2	29.9	4328	23	ABL27278	Drosophila melanog
C 15	30.2	29.9	7514	23	ABL04220	Drosophila melanog
C 16	30	29.7	5863	23	ABL04354	Drosophila melanog
C 17	29.8	29.5	17979	22	AAK64964	Human immune/haema
C 18	29.8	29.5	17979	22	AAK71684	Human immune/haema
C 19	29.4	29.1	580	22	AAK81101	Human polynucleoti
C 20	29.4	29.1	580	22	AAK36027	cDNA sequence #418
C 21	29.4	29.1	788	22	AAI95236	Human neuroblastom
C 22	29.4	29.1	788	22	AAI98133	Human neuroblastom
C 23	29.4	29.1	3707	22	AAO08367	Human secreted pro
C 24	29.4	29.1	429	24	ABT03068	Human breast speci
C 25	29	28.9	5298	24	ABL33635	Human immune syste
C 26	28.6	28.3	2227	24	ABK14977	Human cDNA encodin
C 27	28.2	28.0	6379	22	AAK46347	Tumour suppressor
C 28	28.6	28.3	415	22	AAF67314	Novel human polynu
C 29	28.6	28.3	6686	24	ABL32460	Human immune syste
C 30	28.6	28.3	75859	24	ABK65261	Human genomic DNA
C 31	28.6	28.3	147724	24	ABK83566	Human cDNA differe
C 32	28.4	28.1	662	22	AAI96700	Human neuroblastom
C 33	28.4	28.1	379	24	ABL61747	Colon adenocarcino
C 34	28.2	27.9	439	23	ABV57525	Human prostate exp
C 35	28.2	27.9	1212	21	AAZ65332	Human secreted pro
C 36	28.2	27.9	1608	20	AAZ00815	Human immune syste
C 37	28.2	27.9	10020	24	ABL34292	Human secreted pro
C 38	28.2	27.9	12356	22	AAK46510	Tumour suppressor
C 39	28	27.7	195	22	AAK17824	Human secreted pro
C 40	28	27.7	380	22	AAK57410	Human foetal liver
C 41	28	27.7	380	22	AAK05446	Human bone expres
C 42	28	27.7	380	22	AAK31045	Human bone marrow
C 43	28	27.7	380	22	AAI36955	Probe #5611 used t
C 44	28	27.7	380	24	ABSO5795	Human genome-deriv
C 45	28	27.7	5000	22	AAS14506	Human GSTT1 deleti

ALIGNMENTS

RESULT 1
AAD19456
ID: AAD19456 standard: DNA; 1256 BP.

XX AC AAD19456;

XX DI 18-DEC-2001 (First entry)

XX DE Human fit-1 gene DNA fragment #5.

XX KW Human; fit-1 gene; cytosolic; haemostatic; gene mapping; gynaecological;

KW ophthalmological; rheumatoid arthritis; endometriosis; arthrogenic disease;

KW diabetic retinopathy; psoriasis; VEGFR-1; drug therapy; pharmacogenetic;

KW vascular endothelial growth factor; single nucleotide polymorphism; SNP;

XX KW Cancer; medication; ds.

OS Homo sapiens.

Key	Location/Qualifiers
FI intron	1..579
FI exon	580..664
FI intron	/*tag= a
FI intron	665..1256
FI intron	/*tag= b
FI variation	/*tag= c
FI variation	replace (696, C)
FI variation	/*tag= d
FI variation	/*standard_name= "Single nucleotide polymorphism (SNP)"

PN EP1130123-A2.

XX 05-SEP-2001.

XX WO200171042-A2.
 XX
 XX PD 27-SEP-2001.
 XX
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX
 XX PA (PEKE) PE CORP NY.
 XX
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX DR WPI; 2001-656860/75.
 XX
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Claim 1; SEQ ID NO 3307; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (AB57737-AB572072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX SQ Sequence 4328 BP; 1144 A; 996 C; 1077 G; 1111 T; 0 other;
 Query Match 29.9%; Score 30.2; DB 23; Length 4328;
 Best Local Similarity 62.7%; Pred. No. 18;
 Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 20 CTGACTCATTAGAGTGGATGCGCTTTTCCTGCTTGATATTCITTCIGTTCATTA 79
 DB 3550 CTGCTAAAGAGTGTAGCCCTTACTTTTTTGGAAAGTTCITTCATTA 3609
 QY 80 GATATGTAATGCTTT 94
 DB 3610 ATAATATTAATTTT 3624
 SOLT 15
 ABL04220
 ID ABL04220 standard; cDNA; 7914 BP.
 XX
 XX AC ABL04220;
 XX
 XX DT 26-MAR-2002 (first entry)
 XX
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 7142.
 XX
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical; gene; ss.
 XX
 XX OS Drosophila melanogaster.
 XX
 XX PN WO200171042-A2.
 XX
 XX PD 27-SEP-2001.
 XX
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX

PA (PEKE) PE CORP NY.
 XX
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX DR WPI; 2001-656860/75.
 XX
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Claim 1; SEQ ID NO 7142; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (AB57737-AB572072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX SQ Sequence 7514 BP; 1679 A; 2004 C; 1981 G; 2250 I; 0 other;
 Query Match 29.9%; Score 30.2; DB 23; Length 7514;
 Best Local Similarity 62.7%; Pred. No. 21;
 Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 20 CTGACTCATTAGAGTGGATGCGCTTTTCCTGCTTGATATTCITTCIGTTCATTA 79
 DB 306 CTGCTAAAGAGTGTAGCCCTTACTTTTTTGGAAAGTTCITTCATTA 965
 QY 80 GATATGTAATGCTTT 94
 DB 566 ATAATATTAATTTT 980
 Search completed: December 8, 2002, 06:02:36
 Job time : 134.05 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

R.n on: December 8, 2002, 05:16:46 ; Search time 801.97 Seconds
(without alignments)
3665.203 Million cell updates/sec

Title: US-09-778-900A-5_COPY_500_600

Perfect score: 101

Sequence: 1 atattgtccgtgtgtcc.....tatgttaaatgcttccaagtt 101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

tal number of hits satisfying chosen parameters: 4109250

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hig.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_lm.*
- 20: em_lm.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hig_hum.*
- 31: em_hig_inv.*
- 32: em_hig_other.*
- 33: em_hig_mus.*
- 34: em_hig_pln.*
- 35: em_hig_ror.*
- 36: em_hig_mam.*
- 37: em_hig_vit.*
- 38: em_sy.*
- 39: em_higto_hum.*
- 40: em_higto_mus.*
- 41: em_higto_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	99.4	98.4	104519	9	AL138712 Human DNA
C 2	99.4	98.4	124430	2	AC016311 Homo sapi
C 3	97.2	56.6	101587	2	AC128808 Rattus no
C 4	96.6	36.2	170443	2	AC116776 Mus muscu
C 5	96.2	35.8	96559	9	AC113610 Homo sapi
C 6	95.4	35.0	172076	2	AC121895 Mus muscu
C 7	95.4	35.0	25813	2	AC094998 Rattus no
C 8	95.2	34.9	18110	9	AC098158 Homo sapi
C 9	95.2	34.9	15223	2	AC124802 Homo sapi
C 10	95.2	34.9	106010	2	AC129927 Homo sapi
C 11	95.2	34.7	125971	2	AC119379 Rattus no
C 12	94.2	33.9	8553	2	AC091068 Homo sapi
C 13	94.2	33.7	19378	2	AC068780 Homo sapi
C 14	93.8	33.5	19207	2	AC128050 Rattus no
C 15	93.8	33.5	125733	2	AC126108 Rattus no
C 16	93.8	33.5	192003	2	AC128520 Rattus no
C 17	93.4	33.1	61096	2	AC113537 Mus muscu
C 18	93.4	33.1	15480	2	AC068074 Homo sapi
C 19	93.4	33.1	173818	9	AC098043 Homo sapi
C 20	93.4	33.1	135536	10	AC094416 Mus muscu
C 21	93.4	33.1	195007	2	AC124127 Mus muscu
C 22	93.4	33.1	235628	2	AC102099 Mus muscu
C 23	93.2	32.9	195230	2	AC107661 Mus muscu
C 24	93.2	32.9	213969	2	AC108448 Homo sapi
C 25	93.2	32.9	226354	2	HSA242P20 Homo sapi
C 26	93.2	32.7	14496	10	AL645975 House DNA
C 27	93.2	32.7	49433	9	AL157707 Human DNA
C 28	93.2	32.7	67523	2	AC104356 Homo sapi
C 29	92.6	32.3	7361	2	AC073512 Homo sapi
C 30	92.6	32.3	147634	2	AC123965 Macaca mu
C 31	92.6	32.3	153452	2	AC044895 Homo sapi
C 32	92.6	32.3	156929	9	AC084730 Papio cyn
C 33	92.6	32.3	172913	2	AC124158 Macaca mu
C 34	92.6	32.3	178572	2	AC107346 Rattus no
C 35	92.6	32.3	191264	2	AC096096 Rattus no
C 36	92.6	32.3	194462	2	AC101700 Mus muscu
C 37	92.6	32.3	201332	9	AC091194 Homo sapi
C 38	92.6	32.3	202337	2	AC074394 Homo sapi
C 39	92.4	32.1	166188	9	AC011450 Homo sapi
C 40	92.4	32.1	166448	2	AL732545 Mus muscu
C 41	92.4	32.1	133042	9	AC011354 Homo sapi
C 42	92.4	32.1	135245	9	AC073907 Homo sapi
C 43	92.4	32.1	137848	9	AC097479 Homo sapi
C 44	92.4	32.1	151276	2	AC106956 Rattus no
C 45	92.4	32.1	155471	2	AC015713 Homo sapi

ALIGNMENTS

RESULT 1
AL138712/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-502P18 on chromosome 13.
complete sequence.
ACCESSION AL138712
VERSION AL138712.19 GI:15131448
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164519)
AUTHORS Bates.K.
TITLE Direct Submission

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy, Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinn, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, J., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Haviak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, E.,
 Homs, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, E.,
 Jacobson, B., Jia, T., Johnson, R., Jolivet, S., Joudan, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Krawovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, H.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louis, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, J., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nockenkwo, S., Oguh, M., Okunuga, C.,
 Oquendo, N., Oriado, R., Pace, A., Payton, B., Peery, J., Perez, J.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Rea, J.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, J.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
 Sodergren, E., Sosa, T., Sparks, A., Stanley, H., Store, E.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tareisa, K., Taz, B.,
 Tansey, J., Taylor, C., Taylor, J., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalobos, D., Vinson, R., Wang, S.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gibbs, R.

Direct Submission
 Unpublished
 2 (bases 1 to 101987)
 Worley, K.C.

Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center -----
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information -----
 Center project name: XMMN
 Center clone name: CH230-69N22
 ----- Summary Statistics -----
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 55655 bases at least Q40
 Consensus quality: 59929 bases at least Q30
 Consensus quality: 62415 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length.
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 59 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1002: contig of 1002 bp in length
 * 1003 1102: gap of unknown length
 * 1103 2129: contig of 1027 bp in length
 * 2130 2229: gap of unknown length
 * 2230 3237: contig of 1008 bp in length
 * 3238 3337: gap of unknown length

3338 4734: contig of 1447 bp in length
 * 4735 4884: gap of unknown length
 * 4885 6614: contig of 1730 bp in length
 * 6615 6715: gap of unknown length
 * 6716 8257: contig of 1543 bp in length
 * 8258 9401: contig of unknown length
 * 9402 9501: contig of 1044 bp in length
 * 9502 10661: contig of unknown length
 * 10662 10761: gap of unknown length
 * 10762 11930: contig of 1169 bp in length
 * 11931 12030: gap of unknown length
 * 12031 13552: contig of 1522 bp in length
 * 13553 13632: gap of unknown length
 * 13633 15262: contig of 1610 bp in length
 * 15263 15362: gap of unknown length
 * 15363 16503: contig of 1141 bp in length
 * 16504 17584: gap of unknown length
 * 17585 18054: contig of 1361 bp in length
 * 18055 19537: contig of 1473 bp in length
 * 19538 19637: gap of unknown length
 * 19638 21165: contig of 1528 bp in length
 * 21166 22824: contig of 1559 bp in length
 * 22825 23924: gap of unknown length
 * 23925 23955: contig of 1041 bp in length
 * 23956 24065: gap of unknown length
 * 24066 25710: contig of 1645 bp in length
 * 25711 27147: contig of 1337 bp in length
 * 27148 27247: gap of unknown length
 * 27248 28560: contig of 1613 bp in length
 * 28561 28961: gap of unknown length
 * 28962 30063: contig of 1003 bp in length
 * 30064 31123: contig of 1060 bp in length
 * 31124 31223: gap of unknown length
 * 31224 32567: contig of 1644 bp in length
 * 32568 32957: gap of unknown length
 * 32958 34070: contig of 1103 bp in length
 * 34071 34170: gap of unknown length
 * 34171 36370: contig of 2200 bp in length
 * 36371 37801: contig of 1331 bp in length
 * 37802 37901: gap of unknown length
 * 37902 39087: contig of 1186 bp in length
 * 39088 39187: gap of unknown length
 * 39188 40712: contig of 1525 bp in length
 * 40713 42242: gap of unknown length
 * 42243 43342: gap of unknown length
 * 43343 43567: contig of 1225 bp in length
 * 43568 43667: gap of unknown length
 * 43668 44985: contig of 1318 bp in length
 * 44986 45085: gap of unknown length
 * 45086 46200: contig of 1115 bp in length
 * 46201 46300: gap of unknown length
 * 46301 47472: contig of 1172 bp in length
 * 47473 47572: gap of unknown length
 * 47573 49438: contig of 1856 bp in length
 * 49439 49528: gap of unknown length
 * 49529 51428: contig of 1900 bp in length
 * 51429 51528: gap of unknown length
 * 51529 52659: contig of 1171 bp in length
 * 52700 52799: gap of unknown length
 * 52800 54476: contig of 1677 bp in length
 * 54477 54576: gap of unknown length
 * 54577 55912: contig of 1336 bp in length
 * 55913 56012: gap of unknown length
 * 56013 57709: contig of 1697 bp in length
 * 57710 57809: gap of unknown length
 * 57811 59785: contig of 1980 bp in length

```

* 1 953: contig of 953 bp in length.
* 954 1053: gap of 100 bp
* 1054 7617: contig of 8364 bp in length.
* 7618 16955: contig of 9238 bp in length.
* 16956 17055: gap of 100 bp
* 17056 28121: contig of 11066 bp in length.
* 28122 28221: gap of 100 bp
* 28222 60294: contig of 32073 bp in length.
* 60295 60394: gap of 100 bp
* 60395 95569: contig of 35175 bp in length.
* 95570 95669: gap of 100 bp
* 95670 131654: contig of 35985 bp in length.
* 131655 131754: gap of 100 bp
* 131755 170443: contig of 36689 bp in length.
FEATURES
    source
        1..170443
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="RP24-47245"
            /clone_lib="RC1-24 Male Mouse BAC"
        1..953
            /note="assembly_fragment"
        1054..7617
            /note="assembly_fragment"
        7718..16955
            /note="assembly_fragment"
        17056..28121
            /note="assembly_fragment"
        28222..60294
            /note="assembly_fragment"
        60395..95569
            /note="assembly_fragment"
        95670..131654
            /note="assembly_fragment"
        131755..170443
            /note="assembly_fragment"
BASE COUNT 54653 a 32488 c 32968 g 49630 t 704 others
ORIGIN
Query Match 36.2%; Score 36.6; DB 2; Length 170443;
Best Local Similarity 62.68; Pred. No. 3.5;
Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 10 CTGTGTCACCATTTTCATGATTAACTGAGGGCTTACCTGTCGTTTCGTTGATTAATTCCTGT 69
DB 62057 CTGTGTCACCATTTTCATGATTAACTGAGGGCTTACCTGTCGTTTCGTTGATTAATTCCTGT 61998
Y 70 TTCTCATGATATGTAAGGCTTTCAAGT 100
DB 61997 TTCTCATGATATGTAAGGCTTTCAAGT 61967

RESULT 5
AC113610/C
LOCUS AC113610 96559 bp DNA linear PRI 14-JUL-2002
DEFINITION Homo sapiens BAC clone RP11-373H2 from 2, complete sequence.
ACCESSION AC113610
VERSION AC113610.4 GI:21694035
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 96559)
    AUTHORS Sulston, J.E. and Waterston, R.
    TITLE Toward a complete human genome sequence
    JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
    MEDLINE 99063792
    REFERENCE 2 (bases 1 to 96559)
    AUTHORS Abbott, S., Haakenson, W. and Creason, K.
    TITLE The sequence of Homo sapiens BAC clone RP11-373H2
    JOURNAL Unpublished (2001)

```

REFERENCE

- 3 (cases 1 to 96559)
Waterston, R.H.
Direct Submission
Submitted (04-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
- 4 (cases 1 to 96559)
Waterston, R.H.
Direct Submission
Submitted (30-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
- 5 (cases 1 to 96559)
Waterston, R.H.
Direct Submission
Submitted (04-JUL-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
- 6 (cases 1 to 96559)
Waterston, R.
Direct Submission
Submitted (14-JUL-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 4, 2002 this sequence version replaced gi:20345525.
- Center: Washington University Genome Sequencing Center
Center Code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@wustl.edu
----- Summary Statistics
Center project name: H_NH0373H02

NOTES: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John W. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tator, C.M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Plaster de Jong and co-workers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-62516, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-373H2; actual end is at base position 63477 of RP11-62516.

The region from 20416 to 20419 is covered only by a per product from clone DNA.

FEATURES
Location/Qualifiers
1..96559
/organism="Homo sapiens"
/db_xref="taxon:9606"

Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funk, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, S.,
Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kang, I.,
Karatas, A., Lehocsky, J., Lieu, C., Locke, K., Macdonald, P.,
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Melidrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Rile, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Teshaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J., and Zody, M.

Direct Submission

Submitted (28-JUL-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 146110)

Authors

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginder, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Karatas, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Melidrim, J., Meneus, I.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, A.,
Norbu, C., Norman, C. H., O'Connor, P., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talmes, J., Teshaye, S., Theodorou, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 146110)

Authors

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginder, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Karatas, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Melidrim, J., Meneus, I.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, A.,
Norbu, C., Norman, C. H., O'Connor, P., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talmes, J., Teshaye, S., Theodorou, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 24, 2002 this sequence version replaced g118640683.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research.
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L552
Center clone name: 42_F_20

FEATURES	Location/Qualifiers
Source	1..148110
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="17"
	/map="17"
	/clone="RP11-42F20"
	/clone_lib="RP11 Human Male BAC"
repeat_region	344..1086
	/rpt_family="Charliela"
	/complement(1201..1280)
repeat_region	1281..1389
	/rpt_family="MER2"
repeat_region	1390..1717
	/rpt_family="AluX"
	/complement(1590..1717)
repeat_region	1721..2327
	/rpt_family="MER2"
repeat_region	2328..2901
	/rpt_family="Charliela"
	/complement(2593..2901)
repeat_region	2902..3536
	/rpt_family="AluX"
	/complement(3238..3536)
repeat_region	3537..4392
	/rpt_family="AluX"
	/complement(3546..4392)
repeat_region	4401..4632
	/rpt_family="SVA"
repeat_region	4633..5033
	/rpt_family="L1MC4a"
repeat_region	5034..5053
	/rpt_family="AluX"
repeat_region	5054..5193
	/rpt_family="(TA)n"
repeat_region	5194..5328
	/rpt_family="L2"
repeat_region	5329..5465
	/rpt_family="L1MC4a"
repeat_region	5466..5560
	/rpt_family="FLAM_C"
repeat_region	5561..5868
	/rpt_family="L1MC4a"
repeat_region	5869..6555
	/rpt_family="AluX"
repeat_region	6556..7172
	/rpt_family="L1R33A"
repeat_region	7173..7941
	/rpt_family="AluX"
repeat_region	7942..8242
	/rpt_family="L1PA13"
repeat_region	8243..8533
	/rpt_family="MER58B"
repeat_region	8534..8808
	/rpt_family="AluX"
repeat_region	8809..9360
	/rpt_family="MER58B"
repeat_region	9361..11422
	/rpt_family="L1PA13"
repeat_region	11423..11721
	/rpt_family="AluX"
repeat_region	11722..12006
	/rpt_family="AluX"
repeat_region	12007..13948
	/rpt_family="AluX"
repeat_region	13949..14228
	/rpt_family="L1MC4a"
repeat_region	14229..14625
	/rpt_family="AluX"
repeat_region	14626..14673
	/rpt_family="AT-rich"
repeat_region	14674..15355
	/rpt_family="(T)n"

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

CM: nucleic - nucleic search, using sw model

Run on: December 8, 2002, 05:54:07 ; Search time 2336.87 Seconds
(without alignments)
2086.058 Million cell updates/sec

File: US-09-778-900A-3_COPY_200_500

Perfect score: 301
Sequence: 1 ggcactcgggtgaatg.....taacaaaaactcttaagat 301

Scoring table:

IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pla:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	112.8	37.5	936	13	BI762725 603048490
2	100	33.2	666	9	AL042601 DXZ2p434L
3	95.8	31.8	561	9	AA066758
4	95.8	31.8	855	9	AI326556 mm09f03.Y
5	95.8	31.8	3976	11	AK004721
6	82.8	27.5	978	17	CNS025AS

C	7	31.8	22.2	888	17	CNS04EJU
C	8	77.8	23.8	623	12	EF087145
C	9	77	25.6	868	12	BQ461750
C	10	76.2	25.3	1067	14	BM925760
C	11	75.2	25.0	689	13	BJ075253
C	12	74.6	24.8	463	10	AM681286
C	13	74.6	24.8	656	10	B5577273
C	14	68.8	22.9	609	13	BJ492545
C	15	68	22.6	508	9	AI641359
C	16	67.4	22.4	435	9	AI558355
C	17	65.8	21.9	557	9	AI558605
C	18	65.4	21.7	553	12	BF091324
C	19	65.4	21.7	553	12	BF091324
C	20	65.4	21.7	553	12	BF091324
C	21	65.4	21.7	553	12	BF091324
C	22	65.4	21.7	553	12	BF091324
C	23	65.2	21.6	899	12	B5863321
C	24	65	21.6	899	12	B5863321
C	25	64.8	21.5	229	12	BF444742
C	26	64.8	21.5	481	9	AI092836
C	27	64.8	21.5	481	9	AI092836
C	28	64.2	21.3	471	12	BG011715
C	29	63.8	21.2	603	12	BG41741
C	30	63.8	21.2	656	12	BG713006
C	31	63.8	21.2	656	12	BG713006
C	32	63.6	21.1	286	14	B267483
C	33	63.6	21.1	466	10	B482723
C	34	63.6	21.1	433	14	BQ45275
C	35	63.6	21.1	446	12	BF593396
C	36	63.6	21.1	473	14	R0475
C	37	63.6	21.1	488	13	BI033051
C	38	63.6	21.1	514	10	AV666876
C	39	63.6	21.1	536	12	B808865
C	40	63.6	21.1	536	13	BM030207
C	41	63.6	21.1	541	12	B5757540
C	42	63.6	21.1	557	13	BI540000
C	43	63.6	21.1	567	10	AV614826
C	44	63.6	21.1	571	13	BI755686
C	45	63.6	21.1	500	10	AV601620

ALIGNMENTS

RESULT 1
BI762725

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI762725 603048490f1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:518659 5',
936 bp mRNA linear EST 25-SEP-2001

mRNA sequence.

BI762725

BI762725.1 GI:15754291

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 936)

NIH-MGC <http://mgc.ncl.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-f@mail.nih.gov

Tissue Procurement: life technologies, inc.

cDNA library preparation: life technologies, inc.

cDNA library arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LIML at:

<http://image.llnl.gov>

Plate: LLNL1471 row: c column: 24

High quality sequence stop: 757.

Location/Qualifiers

1..936

FEATURES

SOURCE

GenCore version 5.1.3
 Copyright (c) 1993 - 2002 Compugen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: December 8, 2002, 07:10:23 ; Search time 75.3682 Secs:ds
 (without alignments)
 1478.802 Million cell updates/sec
 Title: US-09-778-900A-3_COPY_200_500
 Perfect score: 301
 Sequence: 1 ggacactcgggttgatgtg.....taacaaaaaacctcttaagt 3..i
 Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0
 Searched: 350425 seqs, 15496369 residues
 tal number of hits satisfying chosen parameters: 700850
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published_Applications_NA:*
 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.6	24.8	5406	10	US-09-919-408-5
2	74.6	24.8	5406	10	US-09-872-136-5
3	74.6	24.8	5470	10	US-09-766-678-1
4	67.2	22.3	3120	10	US-09-982-610-22
5	67.2	22.3	3501	10	US-09-919-408-3
6	67.2	22.3	3501	10	US-09-872-136-3
7	65.6	21.8	4425	10	US-09-982-610-31
8	65.6	21.8	6827	10	US-09-982-610-17
9	65.6	21.8	9108	10	US-09-982-610-45
10	65.4	21.7	3321	10	US-09-866-510-21
11	65.2	21.7	332	10	US-09-960-352-10875
12	65.2	21.7	355	10	US-09-960-352-1177
13	64.8	21.5	3270	10	US-09-866-510-1
14	64.8	21.5	3270	10	US-09-866-510-3
15	64.8	21.5	3270	10	US-09-866-510-5
16	64.8	21.5	3270	10	US-09-866-510-7
17	64.8	21.5	3270	10	US-09-866-510-9
18	64.8	21.5	4054	9	US-09-955-363-35
19	64.8	21.5	6378	10	US-09-919-497-40

20	64.8	21.5	6412	10	US-09-769-987-1	Sequence 1, Appli
21	63.8	21.2	3321	10	US-09-866-510-13	Sequence 13, Appli
22	63.8	21.2	3321	10	US-09-866-510-15	Sequence 15, Appli
23	63.8	21.2	3321	10	US-09-866-510-17	Sequence 17, Appli
24	63.8	21.2	3321	10	US-09-866-510-19	Sequence 19, Appli
25	63.8	21.2	4465	9	US-09-955-363-1	Sequence 1, Appli
26	63.6	21.1	366	10	US-09-960-352-10578	Sequence 10878, A
27	63.6	21.1	391	10	US-09-960-352-2603	Sequence 2603, Ap
28	63.6	21.1	392	10	US-09-960-352-4293	Sequence 4293, Ap
29	63.6	21.1	402	10	US-09-960-352-14442	Sequence 14442, A
30	63.6	21.1	415	10	US-09-960-352-10423	Sequence 10423, A
31	63.6	21.1	3454	12	US-10-044-050-48	Sequence 48, Appli
32	63.6	21.1	3726	10	US-09-925-302-271	Sequence 271, App
33	60.4	20.1	410	10	US-09-960-352-4303	Sequence 4303, Ap
34	58.8	19.5	4268	10	US-09-954-456-293	Sequence 293, App
35	58.8	19.5	4268	10	US-09-954-456-1599	Sequence 1599, Ap
36	58.6	19.5	3992	10	US-09-944-807-9	Sequence 9, Appli
37	57.2	19.0	2184	10	US-09-728-952-82	Sequence 82, Appli
38	56.4	18.7	5084	9	US-09-954-531-1383	Sequence 1383, Ap
39	56.4	18.7	5084	10	US-09-967-7684-277	Sequence 277, App
40	55.6	18.5	434	10	US-09-960-352-5227	Sequence 5227, Ap
41	55.6	18.5	3453	10	US-09-919-408-1	Sequence 1, Appli
42	55.6	18.5	3453	10	US-09-872-136-1	Sequence 1, Appli
43	55.2	18.3	422	10	US-09-960-352-6956	Sequence 6956, Ap
44	54.8	18.2	2869	9	US-10-016-283-2	Sequence 2, Appli
45	53	17.6	2782	9	US-10-108-605-250	Sequence 250, App

ALIGNMENTS

RESULT 1
 US-09-919-408-5
 ; Sequence 5, Application US/09919406
 ; Patent No. US2002007207A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemischka, Ihor R.
 ; TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL
 ; RECEPTORS AND THEIR LIGANDS
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Imclone Systems Incorporated
 ; STREET: 180 Varick Street
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10014
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/919,408
 ; FILING DATE: 31-Jul-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/977,451
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: US 07/906,397
 ; FILING DATE: 26-JUN-1992
 ; APPLICATION NUMBER: US PCT/US92/05401
 ; FILING DATE: 26-JUN-1992
 ; APPLICATION NUMBER: TW 81102961
 ; FILING DATE: 15-APR-1992
 ; APPLICATION NUMBER: US PCT/US92/02750
 ; FILING DATE: 02-APR-1992
 ; APPLICATION NUMBER: US 07/813,593
 ; FILING DATE: 24-DEC-1991
 ; APPLICATION NUMBER: US 07/793,065
 ; FILING DATE: 15-NOV-1991
 ; APPLICATION NUMBER: US 07/728,913
 ; FILING DATE: 28-JUN-1991


```

: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
:
: FILE REFERENCE: ERM-104.01
:
: CURRENT APPLICATION NUMBER: US/09/866,510
:
: CURRENT FILING DATE: 2001-05-25
:
: PRIOR APPLICATION NUMBER: 60/250,747
:
: PRIOR FILING DATE: 2000-12-01
:
: PRIOR APPLICATION NUMBER: 60/289,103
:
: PRIOR FILING DATE: 2001-05-07
:
: NUMBER OF SEQ ID NOS: 33
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 3
:
: LENGTH: 3270
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (1)..(3267)
:
: OS: 09-866-510-3

```

Query Match	21.5%	Score 64.8	DB 10	Length 3270
Best Local Similarity	75.0%	Pred. No. 1.7e-10		
Matches 81	Conservative	0	Mismatches 27	Indels 0
QY	74	CTTCTCTGAATGGGTCTCTGAATCTATCTTTGACAAATCTACAGCAACAAAGC	133	
DB	2565	CTGCCCGAAGTGGATGGCTCTGAGACATCTTTGACAACCTCTACACACACTGACT	2628	
QY	134	GACGTGGTGCTTACGGAGTATGCTGTGGGAATCTCTTCCTTAGGT	181	
DB	2625	GAGCTGTGCTTATGGCANCTCTCTCTGGGAGATCTTTTCCCTTGGT	2676	

```

RESULT 15
US-09-866-510-5
; Sequence 5, Application US/09866510
; Patent No. US200201113041
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
; US-09-866-510-5

```

```

Query Match      21.5%; Score 54.8; DB 10; Length 3270;
Best Local Similarity 75.0%; Pred. No. 1.7e-10;
Matches 81; Conservative 0; Mismatches 27; Indels 0; Gaps 0

QY 74 CTTCTCTGGAATGGATGGCTCTCTGATCTATCTATCTGACAAATCTACACCAAGAGC 133
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 CIGCCCGTGAAGTGGATGGCTCTGAGAGCATCTTTGACACCTCTACACCACTGAST 2628

QY 134 GACGCTGTCGCTACCGAGTATGTGCTGTGGGAATCTCTCTCTTAGT 181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 263 GATGCTGTCGCTATGCGATCTCTGCTCTGGAGATCTTTCCCTGTTG 2676

```

Search completed: December 8, 2002, 09:25:43
Job time : 98.3682 secs


```

repeat_region 32425..32610
/Note="MERSA repeat: matches 4..189 of consensus"
repeat_region 32946..33076
/Note="FLAM_A repeat: matches 1..133 of consensus"
repeat_region 33231..33323
/Note="MIR repeat: matches 118..222 of consensus"
repeat_region 33433..33747
/Note="ALUyB8 repeat: matches 1..313 of consensus"
repeat_region 35242..35605
/Note="THELB repeat: matches 1..364 of consensus"
repeat_region 35886..36063
/Note="89 copies 2 mer tt 57% conserved"
repeat_region 36786..36959
/Note="MER20 repeat: matches 2..169 of consensus"
repeat_region 37212..37318
/Note="MIR repeat: matches 131..241 of consensus"
repeat_region 37283..37330
/Note="L2 repeat: matches 2688..2739 of consensus"
repeat_region 37659..37781
/Note="FLAM_C repeat: matches 1..125 of consensus"
repeat_region 38866..39047
/Note="MERSA repeat: matches 1..186 of consensus"
repeat_region 39114..39348
/Note="MIR repeat: matches 4..231 of consensus"
repeat_region 40010..40203
/Note="ALUJ0 repeat: matches 86..274 of consensus"
repeat_region 40286..40581
/Note="ALUSP repeat: matches 1..297 of consensus"
repeat_region 41033..41330
/Note="ALUSX repeat: matches 1..299 of consensus"
repeat_region 43786..43971
/Note="MES5B repeat: matches 1..177 of consensus"
repeat_region 44671..44839
/Note="MERSA repeat: matches 8..189 of consensus"
repeat_region 45327..45508
/Note="MER3 repeat: matches 1..193 of consensus"
repeat_region 46166..46290
/Note="LIMC5 repeat: matches 7796..7925 of consensus"
repeat_region 46331..46510
/Note="MIR repeat: matches 74..256 of consensus"
repeat_region 46598..46767

```

```

Query Match 99.5%; Score 299.4; DB 9; Length 164519;
Best Local Similarity 99.7%; Pred. No. 3.3e-78;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGACACTCGGTTGAATGCTTTGTTTCATGCGCCCTCTCTCTCAAGGCCCTTGGCCCT 60
b 87314 GGACACTCGGTTGAATGCTTTGTTTCATGCGCCCTCTCTCTCAAGGCCCTTGGCCCT 87255

QY 61 CTCTAGACTCGACTCTCTGAATGGATGCTCTGAATCTATCTTTGACAAATCTA 120
Db 87254 CTCTAGACTCGACTCTCTGAATGGATGCTCTGAATCTATCTTTGACAAATCTA 87195

QY 121 CAGCACAGACGACGTGCTCTACGAGTATGCTGGGAATCTTCCTTAGG 180
Db 87194 CAGCACAGACGACGTGCTCTACGAGTATGCTGGGAATCTTCCTTAGG 87135

QY 181 TAAATTTGGAGAGAGAAATCAACAGCCAGCAATAAATGCTGCAATCTTGCTG 240
Db 87134 TAAATTTGGAGAGAGAAATCAACAGCCAGCAATAAATGCTGCAATCTTGCTG 87075

QY 241 AATGTCCTTTGGTGGACGCTTTAGATTAGAACTACTGTACAAAATCTTAAAG 300
Db 87074 AATGTCCTTTGGTGGACGCTTTAGATTAGAACTACTGTACAAAATCTTAAAG 87015

QY 301 T 301
Db 87014 T 87014

```

RESULT 2
AC128808/c

```

LOCUS AC128808 101987 bp DNA linear HTG 24-JUL-2002
DEFINITION Rattus norvegicus clone CH230-69N22; *** SEQUENCING IN PROGRESS
***, 59 unordered pieces.
ACCESSION AC128808
VERSION 1 GI:21952667
KEYWORDS HTG; HGSC-PHASE1.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 101987)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.H., Amarantunga,H.C., Are,J.R., Ayale,M., Banks,T.,
Bartch,J., Benton,J., Blinze,K., Blankenburg,K., Blankin,D.,
Bouch,J., Bowie,S., Brivea,M., Brown,E., Brown,M., Bryant,N.P.,
Buray,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Cannon,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chokhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Deinewatte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,I., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Gatta,N., Gill,R.,
Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Herrandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,I.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nockenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalobos,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleciyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 101987)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAMN
Center clone name: CH230-69N22
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 55655 bases at least Q40

```


NAME: TIMIAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/10140
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-144-602B-4

Query Match 27.5%; Score 27.8; DB 1; Length 5643;
Best Local Similarity 62.0%; Pred. No. 7.6;
Matches 44; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 28 TTAGAGCTGGAGCCCTTTCTGTCGATGATAATCTTCTTCTTCATAGATAGTA 87
3771 TTACGCGGCTCCGCTGTTCTCTTTTGTGCTGTTCTTCATAGCGGGAC 3630

QY 86 AATGCTTTCAA 98
DB 3831 AAGTTTTCOA 3841

RESULT 8
US-09-450-852-3
; Sequence 3, Application US/09450852
; Patent No. 6309860
; GENERAL INFORMATION:
; APPLICANT: Chaudhri, Raju S.K.
; APPLICANT: Dyomin, Vadim
; APPLICANT: Dalla-Favera, Riccardo
; TITLE OF INVENTION: CLONING AND USES OF BCL-8
; FILE REFERENCE: 53828-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/450,852
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2340
; TYPE: DNA
; ORGANISM: human
US-09-450-852-3

Query Match 27.18; Score 27.4; DB 4; Length 2340;
Best Local Similarity 57.6%; Pred. No. 8.2;
Matches 49; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 2 TATGGCGCTGGTGGTCTCTGATCATTAGCTGAGCCCTTTCTGTCCTGATAATT 61
DB 1128 TATTTTGTGGAATTTGCACAGCTATATGCTTATTTGCTTCATATTTTCAATC 1187
QY 62 CTCTCTCTTCTTCATTAGATCT 86
DB 1188 CATGTTTCTCATCTTTTATATT 1212

RESULT 9
US-09-082-310-3/C
; Sequence 3, Application US/09082310
; Patent No. 6056526
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN NUCLEIC ACID METHYLASES
; NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,310
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0320 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTNOT07
CLONE: 2124957
US-09-082-310-3

Query Match 26.5%; Score 26.8; DB 3; Length 1894;
Best Local Similarity 64.5%; Pred. No. 12;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 20 CTGACTCATTAGCTGAGCCCTTTCTGTCCTGATGATAATCTTCTGTCCTGATA 79
DB 1466 CITAAITTCATCTCTGGTGATTTTCTCTCTGGAATCTTTTCTGTCCTGATA 1407
QY 80 GA 61
DB 1406 AA 1405

RESULT 10
US-09-575-205-3/C
; Sequence 3, Application US/09575205
; Patent No. 6436693
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN NUCLEIC ACID METHYLASES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/575,205

```

Query Match      26.1%  Score 26.4;  DB 4;  Length 3573;
Best Local Similarity 69.2%;  Pred. NO. 18;
Matches 36;  Conservative 0;  Mismatches 15;  Indels 0;  Caps 0;

y 37  GATGCGTTCCTTCGTCCTCAATCTTCCTTCCTTCATTAGATATGAA 88
      ||| ||||| ||||| ||| ||||| ||| ||||| ||| |||||
b 2813 GTTGCGCTTTTCTAATCTGTTAAATATTTCTATTTTACCACAGATTTAA 2864

RESULT 12

```

RESULT 13
 US-08-477-451-9/c
 Sequence 9, Application US/08477451
 Patent No. 5928865
 GENERAL INFORMATION:
 APPLICANT: Covacci, Antonello
 TITLE OF INVENTION: Helicobacter Pylori Cagi Region
 NUMBER OF SEQUENCES: 46
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: CA
 COUNTRY: USA
 ZIP: 94608-2916
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release 1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477.451
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: McClung, Barbara G.
 REGISTRATION NUMBER: 33,113
 REFERENCE/DOCKET NUMBER: 0335.002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 510-601-2708
 TELEFAX: 510-655-3542
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5599 base pairs
 TYPE: nucleic acid

RESULT 15
US-08-477-451-25/c


```

|||||
Db 1340 GTTGTCTCTCTGGCTGGAGCGCGAGACGGCGCTCAGAGCGCGCGCGCGCGCGA 1399
|||||
QY 901 ACAGAGAGGACGACCTCTGGCGCGCGGCTCTTTGGCGCGCGG 941
|||||
Db 1400 ACAGAGAGGACGACCTCTGGCGCGCGGCTCTTTGGCGCGCGG 1440

RESULT 2
AAT08550
ID AAT08550 standard; DNA; 840 BP.
AC
XX
AC AAT08550;
XX
DE 18-JUN-1996 (first entry)
XX
DE Promoter of human flt gene encoding tyrosine kinase receptor.
XX
KW tyrosine kinase receptor; flt gene; promoter; neovascularisation;
tissue-specific expression; heterologous protein production; ds.
XX
XX Homo sapiens.
XX
Key Location/Qualifiers
CDS 771..833
/*tag= a
XX
XX JPU7289263-A.
XX
XX PN
XX PD 07-NOV-1995.
XX
XX PF 22-APR-1994; 94JP-0084526.
XX
XX PR 22-APR-1994; 94JP-0084526.
XX
XX PA (DAUC ) DAIICHI PHARM CO LTD.
XX
XX PF; 1996-015268/02.
XX
XX DR P-PSDB; AAR77646.
XX
XX PI Promoter of human flt gene encoding receptor type tyrosine kinase.
XX useful for tissue specific expression of heterologous proteins
XX
XX PS Claim 1; Page 10; lipp; Japanese.
XX
XX The DNA (nucleotides -229 to +8) contains a promoter for the human
XX flt gene encoding receptor type tyrosine kinase. The promoter is
XX useful for the tissue specific expression of heterologous proteins.
XX Vectors contg. the DNA were used to transform vascular endothelial
XX cells, such that a protein of interest could be expressed. The
XX ligand controlling the function of the promoter may be used as a
XX neovascularisation inhibitor.
XX
SQ Sequence 840 BP; 120 A; 269 C; 336 G; 115 T; 0 other;

Query Match 70.98; Score 667; DB 17; Length 840;
Best local Similarity 98.6%; Pred. No. 9.5e-125;
Matches 725; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

QY 208 CCGGACTGCGCCTCAGTCTCCGTGCGAAGACACCGTCGCGGAGCGCGCGCAGCTTCC 267
|||||
Db 1 CCGGACTGCGCCTCAGTCTCCGTGCGAAGACACCGTCGCGGAGCGCGCGCAGCTTCC 60
|||||
QY 268 CTTGGATCGGACTTTCCGCCCTAGGCGCAGCGCGGAGCTTCAGCTTGTCCCTTCC 327
|||||
Db 61 CTTGGATCGGACTTTCCGCCCTAGGCGCAGCGCGGAG-TCAGCCCTTGTCTT-CC 118
|||||
QY 328 CAGTTTCGGCGCGCGCGCGCGAGCTAGTAGCGCGGTGGAGGAGTCTGCAAGATTCC 387
|||||
Db 119 CAGTTTCGGCGCGCGCGCGCGAGCTAGTAGCGCGGTGGAGGAGTCTGCAAGATTCC 178
|||||
QY 388 TGAGCGCGATGGCAGGAGGAGGCGCAAGGCAAGAGGCGCGCGCAGACCCCTTGAC 447
|||||

```

```

Db 179 TGAGCGCGATGGCAGGAGGAGGCGCAAGGCGCAAGAGGCGCGGAG-AAAGACCCCTGAAAC 237
QY 448 CTCGCCGCGCGCGCTCCCGGGCGCGCTGCGCAGCACTCCCGACGCGGCTTCGCGCCG 507
|||||
Db 238 CTCGCCGCGCGCGCTCCCGGGCGCGCTGCGCAGCACTCCCGACGCGGCTTCGCGCCG 297
QY 508 GGGCCACCGCGCGCTGCTGGCGCGCGCGCGCGCTCTCGTAGCGCGAGGAAAGC-GAGCCTGG 566
|||||
Db 298 GGGCCACCGCGCGCTGCTGGCGCGCGCGCGCGCTCTCGTAGCGCGAGGAAAGC-GAGCCTGG 357
QY 567 GAGGAAGAGAGGAGTAGGTGGCGCAGCGGATGAGGGTGGGGGACCCCTTTGAGTCAACA 626
|||||
Db 358 GAGGAAGAGAGGAGTAGGTGGCGCAGCGGATGAGGGTGGGGGACCCCTTTGAGTCAACA 417
QY 627 GAAGGAGGTGCGCGGGTAGGAAGTGGGCTGGGGAAGGTTATAATCCCGCCCGCCCTCG 686
|||||
Db 418 GAAGGAGGTGCGCGGGTAGGAAGTGGGCTGGGGAAGGTTATAATCCCGCCCGCCCTCG 477
QY 687 GCTGCTCTTCATCGAGGTCGCGGGAGGCTCGGAGCGCGCGCGGCTCCGCGGCTCCCTC 746
|||||
Db 478 GCTGCTCTTCATCGAGGTCGCGGGAGGCTCGGAGCGCGCGCGGCTCCCTC 537
QY 747 CTCTCCCGCGCAGCGCGCGGCTCGGAGCGCGGCTCCGCGGCTCCGCGGCTCGGCGCCAG 806
|||||
Db 538 CTCTCCCGCGCAGCGCGCGGCTCGGAGCGCGGCTCCGCGGCTCGGCGCCAG 597
QY 807 CGGCGCGCTGCGCGCGGAGGATACCGGGGAGTGTGTCTCTGCTGGAGCGCGGAG 866
|||||
Db 598 CGAGCG-CTGGCGCGCGAGGATACCGGGGAGTGTGTCTCTGCTGGAGCGCGGAG 656
QY 867 ACGGCGCTCAGGCGCGCGCGCGCGGCGGAGCGGAGAGGAGTCTGCGGCGCGG 926
|||||
Db 657 ACGGCGCTCAGGCGCGCGCGCGCGGCGGCGGCGGCGGAGAGGAGTCTGCGGCGCGG 926
QY 927 GTCCTTGGCGCGCGG 941
|||||
Db 717 GTCCTTGGCGCGCGG 731

RESULT 3
AAS46351
ID AAS46351 standard; DNA; 6316 BP.
XX
AC AAS46351;
XX
XX 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #73.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
cytosine methylation; ds.
XX
XX Homo sapiens.
XX
XX W0200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-602752/68.
XX

```


DL 1144 CGAATAACGACTTAACGACCGAATCGTTAACCGG 1107

RESULT 10

AA561171/C

ID AA561171 standard; DNA: 6316 BP.

XX AA561171;

XX 29-JAN-2002 (first entry)

XX Human gene regulation-associated gene oligonucleotide #126.

XX Human: Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preseclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; cesophagial cancer; ds; tumour;
 immunostimulant; cardiac; antinflammatory; coagulant; antiasthmatic;
 nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.

XX Homo sapiens.

XX W0200177375-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP03968.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPITENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI: 2002-017470/02.

XX New nucleic acid sequences from chemically modified genes associated

XX with gene regulation, useful for analysing cytosine methylations for

XX diagnosis and therapy of diseases e.g. severe combined immunodeficiency

XX disease

XX Claim 1: SEQ ID No 130; 26pp; English.

XX The invention relates to 224 nucleic acid sequences comprising at least
 XX 18 bases of a chemically pretreated gene associated with gene regulation
 XX chemical pretreatment converts cytosine bases unmethylated at the
 XX 5-position to uracil or another base with hybridisation behaviour
 XX dissimilar to cytosine, to enable analysis of cytosine methylations.
 XX The DNA sequences, oligomers (or sets/arrays) and method are
 XX useful in the diagnosis of diseases (or predisposition to diseases)
 XX associated with gene regulation and in therapy of such diseases, by
 XX enabling analysis of the cytosine methylation patterns of such genes,
 XX kits are provided. They are especially useful in diagnosis
 XX and therapy of e.g. severe combined immunodeficiency disease, cardiac
 XX disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 XX asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 XX preseclampsia, graft versus-host disease. The present sequence is a
 XX sequence included in the sequence data for this specification and is
 XX associated with the human gene regulation-associated genes.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 6316 BP; 1474 A; 255 C; 1549 G; 3038 T; 0 other;

XX Query Match

XX Best Local Similarity 54.3%; Score 510.8; DB 24; Length 6316;

XX Pred. No. 1.8e-93;

	Matches	671;	Conservative	0;	Mismatches	267;	Indels	0;	Gaps	0;
QY	2	CGAAAAAAGACAGGACACGCTCCCTGGACCTGAGCTGGTTCGACGCTTCCTCCAAAGG	61							
DB	2044	CGAAAAAAGACAGGACACGCTCCCTGGACCTGAGCTGGTTCGACGCTTCCTCCAAAGG	1985							
QY	62	TCCCAAGCAACGCTCAGTTCCTCCCTGAGCGCTCCAGGTCAGTTCCTTCGCGAGGTC	121							
DB	1984	TACCAACAAACGCTCAGTTCCTCCCTGAGCGCTCCAGGTCAGTTCCTTCGCGAGGTC	1925							
QY	122	TCCGTCGCTTCCTAGACCTCTCGGACACAGTCTGAGGGTTCAGGAGCGGCGACAGCG	181							
DB	1924	TCGATACCTTCCTTAACCTTCGAAACAACTTAAAAAATCAAAAACGCAACACG	1865							
QY	182	CGGAGAGCAGGCAAGGAGAGCGGAGCTGGCTCAGTTCCTTCGTCGCAAGACA	241							
DB	1864	CGAAAAAAGCAACAAAAAACAACACGCAACTAGCTCCTCAATCCTTCGTCGCAAGACA	1805							
QY	242	CGTCCGCGAGCGCGCGGCGGCTTCCTTCGATCGGCTTCGCGCCCTAGCGCGCGG	301							
DB	1804	CGTCCGCAACCGGACCACTTCCTTAATCGACTTTCGCGCCCTTAACCAACG	1745							
QY	302	GGGAGCTTCAGCTTCCTTCCTCCAGTTTCGGGGCGGCGGCGGAGCTGAGTAAGCG	361							
DB	1744	ACGAAACCTCAACCTTATCCCTTCCTCCCAATTCGAACGACCCCAAACTAAATAACCG	1685							
QY	362	GGTGGAGGAGTCTGCAAGGATTCCTGAGCGCGATGGGAGGAGGCGGCAAGGCA	421							
DB	1684	AATAAAAAAATCTACAAAAATTTCTTAAGCGGATTAACAAAAAATAAAAAAACA	1625							
QY	422	GAGGCGCGGAGCAAGACACCTGAACCTTCGCGGGCGGCGCTCCCGGGCGCGCGTGGCA	481							
DB	1624	AAAAAGCGAACAATAACCTTAACGAAACGCGCTCCCGAACCGCTCGGCA	1565							
QY	482	GCACCTCCCGACGGCGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	541							
DB	1564	ACACCTCCCGACGGCGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1505							
QY	542	CGTACCGCGGAGGAGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	601							
DB	1504	CGTACCGCGGAGGAGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1445							
QY	602	GGTGGGCGGCGGCTTCAGCTCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	661							
DB	1444	AATAAAAAAAGGCTTAAGTCCACCAAAAAAATAACCAAAATAAATAAATAAATAA	1385							
QY	662	AGGTATTAATCGGCGGCGGCGGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT	721							
DB	1384	AAATTAATTAATCGGCGGCGGCGGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT	1325							
QY	722	CGCGCGGCGGCGGCGGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT	781							
DB	1324	CGCGCGGCGGCGGCGGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT	1265							
QY	782	TCCGCGGCGGCTTCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	841							
DB	1264	TCCGCGGCGGCTTCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1205							
QY	842	GTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT	901							
DB	1204	ATATCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT	1145							
QY	902	CGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	939							
DB	1144	CGAAAAAAGCAACTCTTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1107							

RESULT 11

ABK31261/C

ID ABK31261 standard; DNA: 6316 BP.

XX ABK31261;

XX AC

XX XX

23-APR-2002 (first entry)
Signal transduction associated gene modified complementary DNA #52.
Human; signal transduction associated gene; cytosine methylation state;
CpG island; signal transduction associated disease; solid tumour; cancer;
antitumour; cytostatic; mutant; ds.
Homo sapiens.
OS Synthetic.
XX WO200200926-A2.
XX 03-JAN-2002.
XX 29-JUN-2001; 2001WO-EP07472.
XX 30-JUN-2000; 2000DE-1032529.
XX 01-SEP-2000; 2000DE-1043826.
XX (BPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
WPI; 2002-147896/19.
XX Oligonucleotide for diagnosis and therapy of diseases associated with
signal transduction e.g. cancer, comprises chemically modified genomic
sequences of genes associated with signal transduction -
Claim 1; SEQ ID No 104; 24pp; English.
The present invention relates to chemically modified DNA sequences of
signal transduction associated genes. The DNA sequences are chemically
modified using a solution of bisulphite, hydrogen sulphite or
disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
for detecting the cytosine methylation state (CpG islands) of these
genes, and a method for the diagnosis and/or therapy of genetic and
epigenetic parameters of genes associated with signal transduction.
The genomic DNA can be obtained from cells or cellular components which
contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
histologic object slides, and all their possible combinations. The
sequences of the invention are useful for the diagnosis and therapy of
diseases associated with signal transduction e.g. solid tumours and
cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
sequences of different genes associated with signal transduction, or
their complementary sequences.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from the
European Patent Office.
Sequence 6316 BP; 1474 A; 255 C; 1549 G; 3038 T; 0 other;
Query Match 54.3%; Score 510.8; DB 24; Length 6316;
Best Local Similarity 71.5%; Pred. No. 1.8e-93;
Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;
QY 2 CGAAAAAGACAGCAGCAGCTCCCTGGGACCTGAGCTGCGGAGCTCCCAAGG 61
DB 2044 CGAAAAAGACAGCAGCAGCTCCCTGGGACCTGAGCTGCGGAGCTCCCAAGG 1985
QY 62 TGGCAAGCAAGCTAGTTCCTCCCTCAGGCGTCCAGGTTTCAGTCCCTTGTGCGGAGGTC 121
DB 1984 TACCAACAAAGCTAGTTCCTCCCTCAGGCGTCCAGGTTTCAGTCCCTTGTGCGGAGGTC 1925
QY 122 TCGGTGCTCTTACGCTCTCGGACAGCTGGAAGGGGTGAGAGCGCGGAGAGCG 181
DB 1924 TCGGTACCTTCTTAACTTCTCGAACAACATCTAAAAAATCAAAAAACGACGAAACAG 1865
QY 182 CGGAAGACAGCAGGAGGAGCAGCGGAGCTGCGCTCAGTCTCCGTCGCAACACA 241

DB 1864 CGAAAAAGACAAACAAAAAACAACCGAAGTACGCTCAATCCTCCGTACCAAAACA 1805
QY 242 CCGTCGCGGAGGCGCGGAGCTTCCCTTGGATCGGACTTCCGCCCTTAGCGGAGG 301
DB 1804 CCGTCGCGAAGAAACGCGACCACTTCCCTTAAATCGAATCTTCCGCCCTTAAACCAAG 1745
QY 302 CGGAGCTTCAGCTTGTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 361
DB 1744 ACAGAACTTCAACCTTATCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1685
QY 362 CGTGGAGGAGCTTCCAGGATTCCTGAGCGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
DB 1684 AATAAAAAATCTACAAAAATTTCTTAACCGGATTAACAAAAAACAACAAACA 1625
QY 422 GAGGCGCGGAGCAAGAGCTTGAACCTTCCCGGCGCGCTCCCGGCGCGCTCCCGGCGCGCTCCCG 481
DB 1624 AAAAAGCGGAGCAAGAGCTTGAACCTTCCCGGCGCGCTCCCGGCGCGCTCCCGGCGCGCTCCCG 1565
QY 482 GCACCTCCCGGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 541
DB 1564 ACACCTCCCGGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1505
QY 542 CGTACCGGAGCAAGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 601
DB 1504 CGTACCGGAGCAAGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1445
QY 602 GTTGGGAGCAAGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 661
DB 1444 AATAAAAAACCTTAACTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1385
QY 662 AGTTTAAATCGCGCGCTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 721
DB 1384 AATATTAATCGCGCGCTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1325
QY 722 CGCGCGGAGCAAGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 781
DB 1324 CGCGCGGAGCAAGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1265
QY 782 TCGGCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 841
DB 1264 TCGGCGGAGCAAGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1205
QY 842 GTTGTCTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 901
DB 1204 ATATCTCTTAACTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1145
QY 902 CGAGGAGGAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 939
DB 1144 CGAAAAAGACAGCTTAAACGCGGAGTTCGTTAAGCGG 1107
RESULT 12
ABQ57064/C
ID ABQ57064 standard; DNA; 6321 BP.
XX AC ABQ57064;
XX DT 28-AUG-2002 (first entry)
XX XX Human angiogenesis associated polynucleotide SEQ ID NO 94.
DE Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW inflammation; rheumatoid arthritis; diabetic retinopathy; astulcerds;
KW macular degeneration; inflammatory bowel disease; Crohn's disease;
KW anti-rheumatic; anti-arthritis; antidiabetic; antipsoriatic;
XX anti-arteriosclerotic; ds.
XX OS Homo sapiens.
XX XX W0200246454-A2.
XX FN 13-JUN-2002.
XX PD 13-JUN-2002.
XX XX

PF 06-DEC-2001; 2001WO-EP14320.
 PR 06-DEC-2000; 2000DE-1061338.
 PA (EPIG-) EPIGENOMICS AG.
 XX Schacht O;
 PT WPI; 2002-500450/53.
 DR New nucleic acid fragments from chemically treated
 XX anglogenesis-associated genes, useful for determining methylation
 PT status, e.g. in diagnosis or treatment of cancer
 XX
 PS Claim 1; SEQ ID NO 94; 41pp + Sequence Listing; German.
 XX
 CC the invention relates to a nucleic acid (I) comprising a segment of 16
 bases of chemically pretreated DNA of anglogenesis-associated genes (II)
 having sequences (AB066971-AB067178) or their complements (I), also
 related oligomers, are used to evaluate the methylation status and/or
 single-nucleotide polymorphisms, in anglogenesis-related genes, for
 diagnosis and treatment of eye diseases, proliferative retinopathy,
 CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
 CC diabetic retinopathy, macular degeneration caused by neovascularisation,
 CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
 CC Crohn's disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 6321 BP: 1474 A; 258 C; 1553 G; 3036 T; 0 other;
 Query Match 54.3%; Score 510.8; DB 24; Length 6321;
 Best Local Similarity 71.5%; Pred. No. 1.8e-33;
 Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 2 CGGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGAGCTTCCTCCAAAGG 61
 DB 2049 CGGAAAAGACAGGACAGCTCCCTAAACCTTAACCTTAATTCGAACTTCCTCCAA 1990
 QY 62 TCCCAAGCAAGGTCAGTCCCTCAGCGCTCAGGCTTCCAGTTCAGTTCCTGCGGAGGTC 121
 DB 1989 TACCAACAAAGGTCATTCCTCCCTCAACGCTCCAAATTCATACCTATACCGAAATC 1930
 QY 122 TCCGCTGCTTCTTACACTTCTCGGACAGCTCTGAAGGGTTCAGGAGCGGGGACAGCG 181
 DB 1929 TCCGATACCTTCTTAACCTTCTGAAACAACTTAATAAATCAAAACAGCAACACAC 1870
 QY 182 CGGGAAGAGCAGGCAAGGAGAGCAGCGGAGCTCGGCTCAGTCTCCGTCGCAAGACA 241
 DB 1869 CGAAAAGACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 1810
 QY 242 CGCTCGGAGCGCGCGGACAGCTCCCTTGGATCGGACTTTCGCGCTTAGGCGCCAGCG 301
 DB 1809 CGCTCGGAGCGCGCGGACAGCTCCCTTGGATCGGACTTTCGCGCTTAGGCGCCAG 1750
 QY 302 CGGAGCTTTCAGCTTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 361
 DB 1749 AGGAACTTCAACCTTATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1690
 QY 362 GGTGAGGAGCTTCTGAGGAGTTCCTGAGGCGATGGGAGGAGGAGGAGGAGGAGGAA 421
 DB 1689 AATAAATAAATATCAAAAATTTCTTAACCGGATAAACAAACAAACAAACAAACAA 1630
 QY 422 GAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
 DB 1629 AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1570
 QY 482 GCACTTCCCAAGCGGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
 DB 1569 ACACCTCCCAAGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1510
 QY 542 CGTAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601

DB 1569 CTTAACCCGCAAAAACAGAACCTTAAATAAAAAAATAAATAAAAAACGATAAAA 1450
 QY 602 GGTGGGGGAC CTTGAGCTCACCAGAGAGGAGGTCGCGGGTASGAGTGGCTGGGAA 661
 DB 1449 AATAAATAAC CTTAACCTCAGCAAAAAAATAACGGAATAAAAAATAAAAAA 1390
 QY 662 AGGTTATAAA CGCCCCCGCCCTCGGCTGCTTTCATCGAGTCCGCGAGGCTGGGAG 721
 DB 1389 AATAATAAAA CGCCCCCGCCCTCGGCTGCTTTCATCGAATCCCGAAAACTCGAA 1330
 QY 722 CGGCGAGCGGAGGAGCTTCTCGGCTCCTCCCGGAGCGGGGGGCTCGGAGCGGGC 781
 DB 1329 CGGCGAAGGAGGAGCTTCTCGGCTCCTCCCGGAGCGGGGGGCTCGGAGCGGGC 1270
 QY 782 TCGGGGCTCGGCTGAGGAGCGGCGGCGGCTTTCGCGAGGATTCGCGGGAAGTG 841
 DB 1269 TCGGAATCTTATACAGCAGCAGGAGGAGGCTTACGAGAAATATCCGAAAAATA 1210
 QY 842 GTGTCTCTTCTGTGAGCGGAGGAGGAGGCTTACGCGCGGCGGCGGCGGCGGAA 901
 DB 1209 ATATCTCTTCTTAACTAAACCGGAGGAGGAGGCTTAAACCGGAGGAGGAGG 1150
 QY 502 CGAGAGGAGGAGGAGCTTCTGCGGCGGCTTCTTTGGCGCG 939
 DB 1149 CGAATAAAGGAGCTTACGAGCGGAGGAGGCTTACCGCG 1112

RESULT 13
 AAV19347
 ID AAV19347 standard; DNA; 2313 BP.
 XX AC AAV19347;
 XX DI 17-AUG-1995 (first entry)
 XX Human soluble VEGF receptor FLT-1 cDNA.
 DE FLT-1; sFLT-1; vascular endothelial growth factor; VEGF; sVEGF-R1;
 KW receptor tyrosine kinase; signal transduction; angiogenesis;
 KW gene therapy; tumour; metastasis; inflammation; psoriasis;
 KW rheumatoid arthritis; haemangioma; diabetic retinopathy;
 KW angiofibroma; macular degeneration; human; ss.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 250..2313
 FI /*tag= a
 FI
 XX W09813071-A1.
 XX PD 02-APR-1998.
 XX PF 24-SEP-1997; 57WO-US17044.
 XX PR 24-SEP-1996; 56US-0026641.
 XX PA (MERI) MERCK & CO INC.
 XX PI Bett AG, Goldman CK, Huckle WR, Kendall RL, Thomas KA;
 XX WIPI: 1998-23042/20.
 XX P-PSDS: AAW4484.
 XX Inhibition of angiogenesis, particularly in tumours - by using DNA
 PI expressing a soluble form of a tyrosine kinase receptor which forms
 PI a dimer with a vascular endothelial growth factor
 XX example 1; Fig 1: 52pp; English.
 XX This cDNA clone codes for human soluble FLT-1 (sFLT-1, see AAW4484),
 CC the tyrosine kinase receptor (TKR) for human vascular endothelial

CC growth factor (VEGF). PCR cloning was used to obtain the flt-1
 CC derived TKR (sVEGF-RI) flt-1 cDNA from a HUVEC cDNA library. The
 CC invention relates to methods of gene therapy for inhibiting
 CC angiogenesis associated with solid tumor growth, tumor metastasis,
 CC psoriasis, rheumatoid arthritis, haemangioma, diabetic retinopathy,
 CC angiofibroma and macular degeneration. For primary tumor growth
 CC and metastasis, this involves transfer of a nucleotide sequence
 CC encoding a soluble form of VEGF receptor to a mammalian (human) host.
 CC preferably using an adenovirus or recombinant plasmid DNA vector.
 CC The transferred nucleotide sequence transcribes mRNA and a soluble
 CC receptor protein which binds to VEGF in extracellular regions
 CC adjacent to the primary tumor and vascular endothelial cells.
 CC Formation of a sVEGF-R/VEGF complex prevents binding of VEGF to the
 CC FLT-1 TKR, antagonising transduction of the normal intracellular
 CC signals associated with vascular endothelial cell-induced tumor
 CC angiogenesis. Expression of a soluble TKR may also impart a
 CC therapeutic effect by binding either with or without VEGFs to form
 CC non-functional heterodimers with full-length VEGF-specific TKRs and
 CC thereby inhibiting the mitogenic and angiogenic activities of
 CC VEGFs.

-J Sequence 2313 BP; 705 A; 537 C; 539 G; 532 T; 0 other;

Query Match 20.7%; Score 194.8; DB 19; Length 2313;
 Best Local Similarity 98.1%; Pred. No. 3.2e-30;
 Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 730 GGGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCTCGGAGCGGCTCGGGGC 789
 Db 1 GGGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCTCGGAGCGGCTCGGGGC 60
 QY 790 TCGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 849
 Db 61 TCGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 118
 QY 850 CTGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 909
 Db 119 CTGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 178
 QY 910 CGGACTCTGCGGCGGCGGCTCTTTGGCGCGGGG 941
 Db 179 CGGACTCTGCGGCGGCGGCTCTTTGGCGCGGGG 210

RESULT 14
 AAV09330
 ID AAV09330 standard; cDNA; 2313 BP.

AAV09330;

13-MAY-1998 (first entry)

Soluble VEGF receptor/inhibitor (sVEGF-RI) encoding cDNA.

Vascular endothelial cell growth factor; VEGF; receptor; inhibitor;
 soluble; treatment; angiogenesis; psoriasis; tumour; arthritis; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 250..2313

/*tag= a

/product= soluble VEGF inhibitor*

U55712380-A.

27-JAN-1998.

21-APR-1994; 94US-0232538.

21-APR-1994; 94US-0232538.

25-MAR-1993; 93US-0038769.

PA (MERI) MERCK & CO INC.
 XX Kendall RL, Thomas KA;
 XX WPI; 1998-12-03/22/11.
 DR P-PSDB; AAW47036.
 XX Nucleic acid encoding soluble form of vascular endothelial cell
 PI growth factor; receptor - and related vector and transformed cells,
 PI expressing soluble inhibitor of VEGF useful for inhibiting
 PI angiogenesis. e.g. for treatment of psoriasis, arthritis, tumours
 PI etc.

Claim 3; Fig 2A-B; 47pp; English.

CC this cDNA encodes a soluble inhibitor of vascular endothelial cell growth
 CC factor (VEGF). The VEGF receptor/inhibitor (sVEGF-RI) represents a
 CC fragment of the VEGF receptor that binds VEGF with high affinity but is
 CC unable to transduce a signal. sVEGF-RI is used to inhibit VEGF activity,
 CC specifically VEGF-induced angiogenesis, particularly for treatment of
 CC psoriasis, rheumatoid arthritis, haemangiomas, angiofibromas, diabetic
 CC retinopathy, neovascular glaucoma or tumour vascularisation. The sVEGF-RI
 CC can be administered topically or intravenously or from slow-release
 CC matrices.

SQ Sequence 2313 BP; 705 A; 537 C; 539 G; 532 T; 0 other;

Query Match 20.7%; Score 194.8; DB 19; Length 2313;
 Best Local Similarity 98.1%; Pred. No. 3.2e-30;
 Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 730 GGGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCTCGGAGCGGCTCGGGGC 789
 Db 1 GGGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCTCGGAGCGGCTCGGGGC 60
 QY 790 TCGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 849
 Db 61 TCGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 118
 QY 850 CTGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 909
 Db 119 CTGGCTGAGCGGCGGAGCGGCGGCTCGGAGCGGCTCGGAGCGGCTCGGGGC 178
 QY 910 CGGACTCTGCGGCGGCGGCTCTTTGGCGCGGGG 941
 Db 179 CGGACTCTGCGGCGGCGGCTCTTTGGCGCGGGG 210

RESULT 15
 AAV01457
 ID AAV01457 standard; cDNA; 2523 BP.

AAV01457;

25-MAR-1998 (first entry)

Human VEGF receptor extracellular domain coding sequence.

Immunoglobulin-like domain; extracellular domain; VEGF; human; primer;
 Vascular endothelial cell growth factor receptor; amplification; PCR;
 truncation; mutant; tumour; vascularisation; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 250..2523

/*tag= a

/product= "VEGF receptor extracellular domain"

/note= "no stop codon is given at the 3' end of the

sequence"

sig_peptide 250..315

/*tag= b

mat_peptide 316..2523

```
FT XX /*tag= c
XX PN JP09255700-A.
XX XX 30-SEP-1997.
XX PD
XX XX 22-MAR-1996; 96JP-0066711.
XX PF
XX XX 22-MAR-1996; 96JP-0066711.
XX PR
XX XX (TOAG ) TOA GOSSEI CHEM IND LTD.
XX *PA
XX XX WPI: 1997-532773/49.
XX DR P-PSDB; AAW36138.
XX DR
XX XX
XX PT Vascular endothelial growth factor binding polypeptide(s) - used in
XX the treatment of solid tumours etc.
XX SQ
XX SQ Sequence 2523 BP; 765 A; 596 C; 589 G; 573 T; 0 other;
XX
XX This is the nucleotide sequence encoding the 1st-7th immunoglobulin-like
XX domains of the extracellular domain of VEGF (vascular endothelial cell
XX growth factor) receptor FLT. The sequence is amplified by the primers
XX AA01458-V01459. The VEGF receptor sequence can be used to generate
XX truncated mutants which comprise the immunoglobulin-like domains 1-3,
XX 1-4, 1-5 or 1-7. The polypeptides can be used in treatment of solid
XX tumours and other diseases accompanied by pathological vascularisation
XX since they inhibit vascularisation by stimulation with VEGF.
XX
XX Query Match 20.7%; Score 194.8; DB 18; Length 2523;
XX Best Local Similarity 98.1%; Pred. No. 3.2e-30;
XX Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 730 GCGGACACTCTCTGGCTCTCCCGGCGAGCGCGGCTCGGAGCGGCTCCGSGGC 789
Db 1 GCGGACACTCTCTGGCTCTCCCGGCGAGCGCGGCTCGGAGCGGCTCCGSGGC 60
QY 790 TCGGTGCGAGCGCGCGGCGCTGCGCGGAGGATTACCGGGAGTGGTGTCTC 849
Db 61 TCGGTGCGAGCGCGCGGCGGCGGAGGATTACCGGGAGTGGTGTCTC 118
QY 850 CTGGTGGAGCGCGAGCGGCGCTCAGGCGCGGCGGCGGCGGCGGCGGCGGAGGA 909
119 CTGGTGGAGCGCGAGCGGCGCTCAGGCGCGGCGGCGGCGGCGGCGGCGGAGGA 178
910 CGGACTCTGGGCGCGGCTTTGGCCCGGG 941
179 CGGACTCTGGGCGCGGCTTTGGCCCGGG 210
```

Search completed: December 8, 2002, 09:36:42
Job time : 242.314 secs


```

Db 5172 ACAGAGACCGAATTTGGCGGTGGGTCTGTTGGTGGCGG 5212

RESULT 10
AX251106/c
LOCUS AX251106 6316 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 74 from Patent WO0168912.
ACCESSION AX251106
VERSION AX251106.1 GI:15984529
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 6316)
AUTHORS Olek.A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with tumor suppressor genes and
        oncogenes
JOURNAL Patent: WO 0168912-A 74 20-SEP-2001;
        Epigenomics AG (DE)
FEATURES
    source
        1..6316
            Location/Qualifiers
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1474 a 255 c 1549 g 3038 t
ORIGIN

Query Match 54.3%; Score 510.8; DB 6; Length 6316;
Best Local Similarity 71.5%; Pred. No. 6.8e-74;
Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 2 CGCAAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCGAGCTTCCTAAAGG 61
DB 2044 CGCAAAAGACGACGACGCTCCCTAAACCTTAACCTTAATTCGCAATTCCTCAAAA 1985
QY 62 TCCGAAGACGCTGAGTTCCTCCCTGAGGCTTCCAGGCTTCAGTTCCTTGGCGAGGTC 121
DB 1984 TACCAAAACAAACGCTCAATTCCTCCCTCAACGCTTCAATACCTTAATACCGAAATC 1925
QY 122 TCCGCTGCTTCCTAGACTTCTCGGACAGCTCTGAAGGGTTCAGGAGCGCGGACAGCG 181
DB 1924 TCCGATACCTTCCTTAACCTTCGAAACAACTTAATAAATCAAAACGACGAAACACG 1865
QY 182 CGGGAAGACGACGACGCTCCCTGAGGCTTCAGGCTTCAGTTCCTTGGCGAGGTC 121
DB 1864 CGAAAAACAAACGACGACGCTCCCTAAACCTTAACCTTAATTCGCAATTCCTCAAAA 1985
QY 242 CGTGGGAGGCGCGCGCTCCCTGAGGCTTCAGGCTTCAGTTCCTTGGCGAGGTC 301
DB 1804 CGGTCGGAAACGCGACCACTTCCTTAATCGAATTCCTCGCCCTTAAACCAACG 1745
QY 302 GCGGAGCTTCAGCTTCGCTTCCTCCAGTTCGCGGGCGCGCTCCCGGCGCGCG?CGCC? 361
DB 1744 ACGAACTTCAACCTTATCCCTTCCCAATTCGAACGACCCCAACCTAATAAAGCG 1685
QY 362 GTGAGGAGGTCTGCAAGGATTCCTGAGGCGATGGGAGGAGGAGGCGCAAGCGCA 421
DB 1684 AATAAAAAAATCTACAAAAATTCCTAAGCGATTAACAAAAAATAAATAAATAAATAA 1625
QY 422 GAGGCGCGGAGCAAGACCTCTGAACCTGCGGGGCGCGCTCCCGGCGCGCG?CGCC? 481
DB 1624 AAAAAACGGAACAAAAACCTTAACCTTACCGAACCGCGCTCCCGAACCCGCG?CGCA 1555
QY 482 GCACCTTCCCGCGCGCTGGCGCGCGGCGACCGGCTTCGTCGGCGCGCGCGCGCTCTG 541
DB 1564 ACACCTTCCCGCGCGCTCGACCGCGACCGCGCTTCGTCGACCGCGCGCGCTCTG 1505
QY 542 CTGACCGGAGGAGCGACCTTCGGGAGGAGAGAGAGGTAGTGGGAGCGCGGAGG 601
DB 1504 CTGACCGGAGGAGCGACCTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 1445
QY 602 GTTGGGAGCCCTTACGCTACCAAGAGGCTCCCGGGGTAGGAGTGGGCTGGGAA 661

```

```

Db 1444 AATAAAAAACCCCTTAACTGTCACCAAAAAAATAACGAAATAAAAAATAAATAA 1385
QY 662 AGGTTAATAATCGCCCGCCCTCGGCTGCTTTCATCGAGGTGCGGGAGGCTCGGAG 721
DB 1384 AATATATAATCGCCCGCCCTCGACTACTTTCATCGAAATCGCGAAAAACTCGAAA 1325
QY 722 CGCGCAGCGGACACTCTCTCTCGCTCCTCCCGGCGAGGCGCGGCTCGGAGCGGCG 781
DB 1324 CGCGCAAAACGACACTCTCTCTGACICTCTCCGCGACGAGGAGGACTCGAACGAC 1265
QY 782 TCCGGGCTCGGGTGCAGGCGCCAGCGGCGCCTTGGCGGAGGATTAACCGGGGAGTG 841
DB 1264 TCCGAAACTCGAATACACGACGCAACGACCCCTAACGAGAAATTTACCGGAAAAA 1205
QY 842 GTTGTCTCTGCTGAGCGGCGGAGACGCGGCTTCAGGCGCGGCGCGGCGCGGCGGAA 901
DB 1204 ATATCTCTTAACCTAAACCGGAAACGAGACGCTCAAAACGCGGAAACCGGACGAG 1145
QY 902 CGAGAGGACGACTCTGGCGCGCGGCTCTTTGGCGCGG 939
DB 1144 CGAAAAACGAACTCTACGACCGAATGTTAACCGG 1107

RESULT 11
AX251869/c
LOCUS AX251869 6316 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 130 from Patent WO0168911.
ACCESSION AX251869
VERSION AX251869.1 GI:15985224
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 6316)
AUTHORS Olek.A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the cell cycle
JOURNAL Patent: WO 0168911-A 130 20-SEP-2001;
        Epigenomics AG (DE)
FEATURES
    source
        1..6316
            Location/Qualifiers
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1474 a 255 c 1549 g 3038 t
ORIGIN

Query Match 54.3%; Score 510.8; DB 6; Length 6316;
Best Local Similarity 71.5%; Pred. No. 6.8e-74;
Matches 671; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 2 CGCAAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCGAGCTTCCTAAAGG 61
DB 2044 CGCAAAAGACGACGACGCTCCCTAAACCTTAACCTTAATTCGCAATTCCTCAAAA 1985
QY 62 TCCGAAGACGCTGAGTTCCTCCCTGAGGCTTCCAGGCTTCAGTTCCTTGGCGAGGTC 121
DB 1984 TACCAAAACAAACGCTCAATTCCTCCCTCAACGCTTCAATACCTTAATACCGAAATC 1925
QY 122 TCCGCTGCTTCCTAGACTTCTCGGACAGCTCTGAAGGGTTCAGGAGCGCGGACAGCG 181
DB 1924 TCCGATACCTTCCTTAACCTTCGAAACAACTTAATAAATCAAAACGACGAAACACG 1865
QY 182 CGGGAAGACGACGACGCTCCCTGAGGCTTCAGGCTTCAGTTCCTTGGCGAGGTC 121
DB 1864 CGAAAAACAAACGACGACGCTCCCTAAACCTTAACCTTAATTCGCAATTCCTCAAAA 1985
QY 242 CGTGGGAGGCGCGCGCTCCCTGAGGCTTCAGGCTTCAGTTCCTTGGCGAGGTC 301
DB 1804 CGGTCGGAAACGCGACCACTTCCTTAATCGAATTCCTCGCCCTTAAACCAACG 1745
QY 302 GCGGAGCTTCAGCTTCGCTTCCTCCAGTTCGCGGGCGCGCTCCCGGCGCGCG?CGCC? 361
DB 1744 ACGAACTTCAACCTTATCCCTTCCCAATTCGAACGACCCCAACCTAATAAAGCG 1685

```

VI 2044 CCGAAAAAACACAGNACACGCTCCCTTAAACCTTAACCTAACCTCCAAAGCTTCCGAAAAA 1985

REFERENCE
1. Olex, A., Piepenbrock, C. and Berlin, K.

1509 CGTACCGCGCAAAAAACGAACTTAAAAATTTATTTAAAAACGAAATAAAA 1450

QY 602 GGTGGGGACCCCTTCAGCTACACAGAGAGGTGCGGGGTAGGAAGTGGCTGGGAA 661
 Db 1449 AATATAAAACCCCTTACGCTACACCAAAAAAATACCGAAATATAAATAAATAA 1390
 QY 662 AGCTTATAAATCGCCCGCCCTCGCTCTTCATCTGAGGTCCGCGGAGGCTCGAG 721
 Db 1389 AATATATAAATCGCCCGCCCTCGCTCTTCATCTGAGGTCCGCGGAGGCTCGAG 1330
 QY 722 CGGCGACGAGGACCTCTCTCTGCTGCTCTCCCGGAGGCGGCGGCTCGAGCGGCG 781
 Db 1329 CGCGCAACCAACCTCTCTCTGCTGCTCTCCCGGAGGCGGCGGCTCGAGCGGCG 1270
 QY 782 TCGGGGCTCGGCTGAGCGGCGGCGGCTCTGCTGCGGAGGATACCGGGGAGTG 841
 Db 1269 TCGGAATCTGAATACAGCAACGAGCGGCTTACGAGCAAAATATCCCGAAAAA 1210
 QY 842 GTTCTCTCTGAGCGGCGGCGGCTTACGAGCGGCGGCGGCGGCGGCGGCGGAA 901
 Db 1209 ATTATCTCTAACTAAACCGGAGGAGGCTCTAAACCGGAGGAGGAGGAGGAA 1150
 QY 902 CGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 939
 Db 1149 CGAAAAACGAACCTCTAACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1112

RESULT 15
 AC128403/c
 LOCUS Rattus norvegicus clone CH230-448N2. *** SEQUENCING IN PROGRESS
 DEFINITION *** 57 unordered pieces.
 ACCESSION AC128409
 VERSION AC128409.1 GI:215909078
 KEYWORDS HTG; HGS_PHASE1.
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 144090)
 Murny D.M., Adams C., Adio-Oduola B., Ali-Adnan, F.P., Allen C.,
 Albrooks S.L., Amaratunga H.C., Are J.R., Ayala M., Banks I.,
 Barbata J., Benton J., Blum K., Blankenburg K., Bonnin D.,
 Bowe J., Bowie S., Brivea M., Brown E., Brown M., Bryant N.P.,
 Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
 Carron T.F., Carter M., Cavazos S.B., Chacko J., Chavez D.,
 Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
 Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
 Delaney K.P., Delgado O., Denn A.L., Ding Y., Dinh H.H.,
 Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,
 Earhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,
 Falls I., Ferraguto D., Flagg N., Ford J., Foster F., Frantz P.,
 Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
 Gorrell J.H., Guerata M., Gunaratne P., Hale S., Hamilton K.,
 Harris C., Harris K., Hart M., Havlak P., Hayes A., Hernandez J.,
 Hernandez O., Hodgson A., Hoque M., Holloway C., Hollins B.,
 Honsi F., Howard S., Huber J., Hulyk S., Hume J., Jackson L.E.,
 Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S.,
 Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovar C.,
 Kravovic J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L.,
 Li J., Li Z., Lichtarge O., Liebo C., Liu J., Liu R., Louised H.,
 Lozano R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,
 Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E.,
 Massey E., Mathew E., McLeod M.P., Meador M., Mei G., Metzger M.,
 Miner G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,
 Moser M., Neal D., Newton J., Nektson N., Nguyen A., Nguyen N.,
 Nguyen N., Nickerson E., Nwokenwo S., Oguh M., Okonofu G.,
 Oreguene N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,
 Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,
 Rives M., Rojas A., Rojibokan I., Rolfe M., Ruiz S., Savary G.,
 Scherer S., Scott G., Shen H., Shoshitari N., Sisson J.,
 Sodergren E., Sonaike I., Sparks A., Stanley H., Stone H.,

Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,
 Tansey J., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,
 Usmani K., Vasquez L., Vera V., Villalon D., Vinson R., Wang Q.,
 Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,
 Williams G., Williamson A., Wlezyk R., Wooden S., Worley K.,
 Xu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,
 Weinstein G. and Gibbs R.
 Direct Submission
 Unpublished
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GYFL
 Center clone name: CH230-448N2
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 101711 bases at least Q40
 Consensus quality: 106747 bases at least Q30
 Consensus quality: 110016 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see <http://www.hgsc.bcm.tmc.edu/docs/genbank.draft.data.html>).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 57 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1636: contig of 1636 bp in length
 1637 1736: gap of unknown length
 1737 3403: contig of 1667 bp in length
 3404 3503: gap of unknown length
 3504 4554: contig of 1051 bp in length
 4555 4654: gap of unknown length
 4655 5864: contig of 1210 bp in length
 5865 5964: gap of unknown length
 5965 7115: contig of 1151 bp in length
 7116 8510: contig of 1295 bp in length
 8511 8610: gap of unknown length
 8611 9955: contig of 1345 bp in length
 9956 10055: gap of unknown length
 10056 11380: contig of 1325 bp in length
 11381 11480: gap of unknown length
 11481 12607: contig of 1127 bp in length
 12608 12707: gap of unknown length
 12708 13863: contig of 1156 bp in length
 13864 13963: gap of unknown length
 13964 15362: contig of 1399 bp in length
 15363 15462: gap of unknown length
 15463 17024: contig of 1562 bp in length
 17025 17124: gap of unknown length
 17125 18879: contig of 1755 bp in length
 18880 18979: gap of unknown length
 18980 20913: contig of 1934 bp in length
 20914 21013: gap of unknown length
 21014 23216: contig of 2203 bp in length
 23217 23316: gap of unknown length
 23317 25554: contig of 2238 bp in length
 25555 25554: gap of unknown length

* 25655 27062: contig of 1408 bp in length
* 27063 27162: gap of unknown length
* 27163 28572: contig of 1410 bp in length
* 28573 28572: gap of unknown length
* 28573 29330: contig of 1238 bp in length
* 29331 30030: gap of unknown length
* 30031 32527: contig of 2497 bp in length
* 32528 32627: gap of unknown length
* 32628 33981: contig of 1254 bp in length
* 33982 33981: gap of unknown length
* 33982 36613: contig of 2632 bp in length
* 36614 36713: gap of unknown length
* 36714 38530: contig of 1817 bp in length
* 38531 38630: gap of unknown length
* 38631 40221: contig of 1591 bp in length
* 40222 40221: gap of unknown length
* 40222 42216: contig of 1895 bp in length
* 42217 42316: gap of unknown length
* 42317 44902: contig of 2586 bp in length
* 44903 45002: gap of unknown length
* 45003 47520: contig of 2518 bp in length
* 47521 47520: gap of unknown length
* 47521 49384: contig of 1754 bp in length
* 49385 49484: gap of unknown length
* 49485 51442: contig of 1958 bp in length
* 51443 51542: gap of unknown length
* 51543 54024: contig of 2482 bp in length
* 54025 54124: gap of unknown length
* 54125 56701: contig of 2577 bp in length
* 56702 56801: gap of unknown length
* 56802 58164: contig of 1363 bp in length
* 58165 58264: gap of unknown length
* 58265 60579: contig of 2415 bp in length
* 60580 60779: gap of unknown length
* 60780 63488: contig of 2709 bp in length
* 63489 63588: gap of unknown length
* 63589 65928: contig of 2340 bp in length
* 65929 66028: gap of unknown length
* 66029 69193: contig of 3165 bp in length
* 69194 69293: gap of unknown length
* 69294 70340: contig of 1047 bp in length
* 70341 70440: gap of unknown length
* 70441 72200: contig of 1760 bp in length
* 72201 72300: gap of unknown length
* 72301 76378: contig of 4078 bp in length
* 76379 76478: gap of unknown length
* 76479 80199: contig of 3721 bp in length
* 80200 80299: gap of unknown length
* 80300 82182: contig of 1883 bp in length
* 82183 82282: gap of unknown length
* 82283 84734: contig of 2452 bp in length
* 84735 84834: gap of unknown length
* 84835 87943: contig of 3109 bp in length
* 87944 88043: gap of unknown length
* 88044 90121: contig of 2078 bp in length
* 90122 90221: gap of unknown length
* 90222 92597: contig of 2376 bp in length
* 92598 92597: gap of unknown length
* 92598 97180: contig of 4483 bp in length
* 97181 97280: gap of unknown length
* 97281 99751: contig of 2471 bp in length
* 99752 99851: gap of unknown length
* 99852 104895: contig of 5044 bp in length
* 104896 104995: gap of unknown length
* 104996 108453: contig of 3468 bp in length
* 108454 108563: gap of unknown length
* 108564 111325: contig of 3362 bp in length
* 111326 112025: gap of unknown length
* 112026 115621: contig of 3596 bp in length
* 115622 115721: gap of unknown length
* 115722 118437: contig of 2716 bp in length
* 118438 118537: gap of unknown length
* 118538 123236: contig of 4599 bp in length

FEATURES

Query Match

40.1%; Score 377; DB 2; Length 144090;

Best Local Similarity 74.0%; Pred. No. 1.6e-52;

Matches 702; Conservative 0; Mismatches 180; Indels 67; Gaps 15;

OY 1 GCCGAAAACACGACGACGCTCCCTCCCTGGACCTAGCTGGTTCAGTGTTCGAGTCTTCCCAAAG 60
DB 60478 SCCGAAAACACACGACGCTCCCTCCCTGGACGCGCGCGCGCGCTCTTCCGAAG 60419
OY 51 GTGCCAAGCAA-CGGTCAGTCCCTCAGCGCTCCAGTTC-AGTGCCTTGTCCCGAGG 118
DB 50418 GTGCTGGCAAGCGCTCCGTTTCCITCGGAGCTCCGCGTCCAAAGTCCCTTAAGCGGAGG 50359
OY 119 GTCTCCGTCGCTTCC-TAGACTTCTCGGACAGCTCTGAAGGGGTCA-GGAGCGCGGGA 176
DB 60358 GTCTCGGCGCTTCTCTCGCGCTCTGGGAACTCTGTGGGGGTCAAGGACCGCGGGA 60299
OY 177 CAGCGCGGGAAGAGGAGGCAAGGGAGACACCGGACTCGCGCTCAGTCTTCCGTGCCAA 236
DB 60298 CCGCTC-----AGGAGAGGCTCGACTCGCGCTCTTCTCTCGGTGCAG 60255
OY 237 GAACACCGTCGCGAGCGCGGCGGCTTCCCTTGGATGGACTTTCGCGCCCTTAGGGCC 296
DB 60254 GGACACCGTCGCGGGGCGCGGCGGCTTCCCTAGGTAAGACTTCCCGCGCGCGGGGCA 60195
OY 297 AGCGCGCGGAGCTTCAGCTTGTCCCTTCCCGAGTTCGCGCGCGCGCGCTCCCGCGCGCG 475
DB 60194 GCGCGGTGCATTT---CGCGGTTCCTCTCTCAGTTTGGGGGTCAACAGAGCTGAGTA 60138
OY 357 AGCGCGGTGGAGGAG-TCGCAAGGATTCCTGAGCGGATGGCGGAGGAGGAGGCA 415
DB 60137 AGCGCGGTGGAGGAGCTGGCAAGGATTCCTGAGAGGATGGGC---AGGAGGGGCGG 60081
OY 415 GGGCAAGAGCGCGGAGCAAGAGCTTGAACCTTGGCGGCGCGCGCTCCCGCGCGCGCG 475
DB 60080 GGGCAGCAGCGGACCGCAGCAGCAGCGCTGACCTGC---GAACCTGCGCGCGGACCGCG 60024
OY 476 TCGCGACACTCTCCACCGCGCGCTC---GGCGCGCGCGCACCGCGCTCTGCGCGCGCG 533
DB 60023 CGCTTAGCGCCACCGCAGCGCGCTCGCGCGCGCGCGCTCGCGCGCGCGCGCG 59964
OY 534 CCGCTCTCCGTAGCGCGCAGGGAAGGAGCGCTGCGGAGGAGAGAGAGGTAGTGGGAGGC 593
DB 59963 CCGCTC-----GGAGGAGAGAGAGGTAGTGGGAGGC 59928
OY 594 GGATGAGGCTGGGGAGCGCTTACGCTACCAAGAGAGGTGCGCGGGTAGGAAGTGGG 653
DB 59927 GGATGAGGCTGGGGAGCGCTTACGCTACCAAGAGAGGTGCGCGGGTAGGAAGTGGG 59868
OY 654 CTGGGAAAGGTTTAAATCGCGCGCGCGCTCGCGCTCTCTCATCGAGGTCCCGCGGAG 713
DB 59867 CTGGGG-AAGGTTTAAATCGCGCGCGCGCTCGCGCTCTCATCGAGGTCTCGGAG 59809
OY 714 GCTGAGCGCGCGCAGCGGACACTCTCTCTCGGCTCTCTCCCG---CGAGCGCGCGCGCT 771
DB 59808 GCTCGAGCGCG-CTGCGGACACTTCCCGGAGGTAGTGGTGTCTGCTGCGACAGCT 59750
OY 772 CGGAGCGGCTCCCGGCTCGGGTTCAGCGGCGCAGCGGCGCTCGCGCGGAGGATACC 831
DB 59749 CGGAGCGGCTCCCGGCTCAAGCGCAGCGCTAGCGGACGCGAGCTGGGAGGATCCC 59690
OY 832 CGGGAAGTGTGTCTCTCTCGGCTCGAGCGCGGAGCGGCGCTCAGGGCGCGCGCG 891
DB 59689 TCGCGCGCGCGCACACCTCTCTCTCGGCTCGAGGCGCGAGAGAGCTCTCGCGCGCGC-----C 59637

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 09:21:42 ; Search time 1219.71 Seconds
(without alignments)
12494.717 Million cell updates/sec

Title: D64016_COPY_500_1440

Perfect score: 941

Sequence: 1 GCCGAAAGACACGGACAC.....GCCGGCTTTGGCCCGGG 941

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154056 seqs, 8097743376 residues

Cal number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_est1:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_Other:**

26: em_gss_pro:**

27: em_gss_rod:**

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243.4	25.9	697	9	AU136034
2	232.4	24.7	775	13	BI488969
3	230.8	24.5	906	13	BI820974
4	230.8	24.5	1171	11	BC029849
5	224.4	23.8	826	9	AL541018
6	223.4	23.7	972	9	AL543052

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

```

RESULT 1
AU136034
LOCUS AU136034 697 bp mRNA linear EST 02-AUG-2002
DEFINITION AU136034 PLACE1 Homo sapiens cDNA clone PLACE1003403 5', mRNA
ACCESSION AU136034
VERSION AU136034.1 GI:10996573
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,,
Yamanoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project, 5' & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

```

7	221.8	23.6	1002	9	AL547815
8	215.4	22.9	662	9	AU136704
9	212.2	22.6	1025	9	AL551816
10	211.4	22.5	638	9	AL548886
11	207.4	22.0	860	9	AL546515
12	206.6	22.0	584	14	BQ271447
13	206.4	21.9	1030	9	AL543703
14	204.4	21.7	740	9	AL549158
15	202	21.5	867	9	AL540382
16	200.4	21.3	755	9	AL548841
17	200.4	21.3	1066	13	BI488800
18	197.8	21.0	1067	13	BM546253
19	196.4	20.9	943	13	BI488796
20	173	18.4	828	9	AL546867
21	158.8	16.9	468	9	AL544285
22	103.4	11.0	989	9	AL544285
23	93.8	10.0	615	14	BM735444
24	89.4	9.5	1160	17	AG043473
25	85.4	9.1	925	17	CNS0091P
26	85.2	9.1	925	17	CNS0091P
27	84	8.9	909	9	AL572624
28	79	8.4	811	13	BI888042
29	78.2	8.3	811	13	BI956687
30	78.2	8.3	1288	14	BQ678719
31	77.4	8.2	1137	12	BG809379
32	77	8.2	949	17	AG043499
33	77	8.2	1052	14	BQ950956
34	76.8	8.2	935	17	CNS006XX
35	76.6	8.1	1165	17	AG030649
36	76.2	8.1	1538	17	AG030607
37	75	8.0	1101	17	AG039543
38	74.8	7.9	1151	14	BQ951254
39	74.4	7.9	810	17	AG126142
40	74.4	7.9	1009	17	CNS010EW
41	74.2	7.9	468	14	BQ608819
42	74.2	7.9	948	17	AG161418
43	74.2	7.9	1143	10	AW731158
44	74	7.9	1128	14	BQ710735
45	73.8	7.8	932	17	CNS0072Q

```

primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."
BASE COUNT      104 a 243 c 324 g 104 t
ORIGIN
Query Match      24.7% Score 232.4; DB 13; Length 775;
Best Local Similarity 99.6%; Pred. No. 5.9e-35;
Matches 233; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 708 CGGAGGCGCGGAGCGGCGGACGCTCTCTCGGCTCTCTCCCGGACGCGCGGC 767
Db 1 CGGAGGCTCGGAGCGCGCGGCGGACGCTCTCTCGGCTCTCTCCCGGACGCGCGGC 60
QY 768 GCCTCGAGCGGGCTCCGGGGCTGGGTGCGAGCGGCAGCGGGCGCTTGCGCGCGGAG 827
Db 51 GGCCTCGAGCGGGCTCCGGGGCTGGGTGCGAGCGGCAGCGGGCGCTTGCGCGCGGAG 120
QY 828 TACCGGGGAAGTGGTGTCTCTCTCGGCTGGAGCGCGGAGAGCGGCGCTCAGGCGCGGG 887
Db 121 TACCGGGGAAGTGGTGTCTCTCTGGCTGGAGCGCGGAGAGCGGCGCTCAGGCGCGGG 180
QY 888 CGGCGCGCGCGCAACGAGAGGACGAGCTCTGGCGCGCGGCTTGCGCGCGGG 941
Db 181 CGGCGCGCGCGCAACGAGAGGACGAGCTCTGGCGCGCGGCTTGCGCGCGGG 234
RESUL 3
BI820974 603035625F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176391
LOCUS mRNA sequence.
DEFINITION
ACCESSION BI820974
VERSION BI820974.1 GI:15932524
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 906)
AUTHORS NIH-National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE MGC clone distribution information can be
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-research.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IMAGE1440 Row: a Column: 08
High quality sequence start: 2
High quality sequence stop: 765.
Location/Qualifiers
1..906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176391"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
Note="Organ: pooled brain, lung, testis; Vector:
pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.

```

Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library.
BASE COUNT 219 a 228 c 254 g 194 t 1 others

ORIGIN

Query Match 24.5%; Score 230.8; DB 13; Length 906;
Best Local Similarity 99.1%; Pred. No. 1.2e-34;
Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 708 CGGAGGCTCGGAGCGCCAGCGGCGACACTCTCTCGGCTCTCTCCCGGCGAGCGGCGC 767

Db 1 CGGAGGCTCGGAGCGCCAGCGGCGACACTCTCTCGGCTCTCTCCCGGCGAGCGGCGC 50

QY 768 GGCTCGAGCGGCTCGGAGCGGCTCGGAGCGGCGGCGGCGGCGGCGGAGAT 827

Db 61 GGCTCGAGCGGCTCGGAGCGGCTCGGAGCGGCGGCGGCGGCGGCGGAGAT 120

QY 828 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGGCGGCGGCGGCGGCGGCGG 887

Db 121 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGGCGGCGGCGGCGGCGGCGG 180

QY 888 CCGCGGCGGCGGAGCGAGCGGACTCTGCGGCGGCGGCGGCGGCGGCGGCGG 941

Db 181 CCGCGGCGGCGGAGCGAGCGGACTCTGCGGCGGCGGCGGCGGCGGCGGCGG 234

RESULT 4
BC029849
LOCUS BC029849 1171 bp mRNA linear HTC 20-MAY-2002
DEFINITION Homo sapiens, clone IMAGE:5176591, mRNA.

ACCESSION BC029849
VERSION BC029849.1 GI:20987442

KEYWORDS HTC.
SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1171)

REFERENCE 1 (bases 1 to 1171)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgpubs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HQSC
Web site: <http://www.hqsc.bcm.tmc.edu/cdna/>
Contact: ang@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hui, K., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLML at: <http://image.llnl.gov>
Series: IRAC Plate: 51 Row: m Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4503748
This clone has the following problem: incomplete processing.

FEATURES Location/Qualifiers
1..1171
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176591"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"
BASE COUNT 340 a 260 c 302 g 269 t

ORIGIN

Query Match 24.5%; Score 230.8; DB 11; Length 1171;
Best Local Similarity 99.1%; Pred. No. 1.2e-34;
Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 708 CGGAGGCTCGGAGCGCCAGCGGCGACACTCTCTCGGCTCTCTCCCGGCGAGCGGCGC 767

Db 1 CGGAGGCTCGGAGCGCCAGCGGCGACACTCTCTCGGCTCTCTCCCGGCGAGCGGCGC 50

QY 768 GGCTCGAGCGGCTCGGAGCGGCTCGGAGCGGCGGCGGCGGCGGAGAT 827

Db 61 GGCTCGAGCGGCTCGGAGCGGCTCGGAGCGGCGGCGGCGGCGGAGAT 120

QY 828 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGGCGGCGGCGGCGGCGGCGG 887

Db 121 TACCGGGGAAGTGTGTCTCTCGGCTCGGAGCGGCGGCGGCGGCGGCGGCGG 180

QY 888 CCGCGGCGGCGGAGCGAGCGGACTCTGCGGCGGCGGCGGCGGCGGCGGCGG 941

Db 181 CCGCGGCGGCGGAGCGAGCGGACTCTGCGGCGGCGGCGGCGGCGGCGGCGG 234

RESULT 5
AL541018
LOCUS AL541018 LRI_FLO02_P11 826 bp mRNA linear EST 15-FEB-2001
DEFINITION AL541018 LRI_FLO02_P11 Homo sapiens cDNA clone CS0DR005Y102 5 prime
mRNA sequence.

ACCESSION AL541018
VERSION AL541018.1 GI:12871688

KEYWORDS EST.
SOURCE Human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 826)

REFERENCE 1 (bases 1 to 826)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91005 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES Location/Qualifiers
1..826
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE005Y102"
/clone_lib="LRI_FLO02_P11"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Peng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville,
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
liang@lifetech.com URL :
<http://fulllength.invitrogen.com>"

BASE COUNT 201 a 218 c 243 g 162 t 2 others

Query Match 23.8%; Score 224.4; DB 9; Length 826;
Best Local Similarity 99.6%; Pred. No. 2e-33;
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 716 TCGGAGCGCGCAGCGGCGACACTCTCTCGGCTCTCTCCCGGCGAGCGGCGGCGCGG 775

Db 1 TCGGAGCGCGCAGCGGCGACACTCTCTCGGCTCTCTCCCGGCGAGCGGCGGCGCGG 50

Db 1886 TTAGAATGATCTTGGAGGACCAAGAGACTAAGTCTGCTTCTGATGAGACAGAG 1845
QY 278 CAGGGAGAAATCTCCAGAGAAGAAATTAATATGAGATCAGGAAGCACCATAACC 337
Db 1946 CCAAGAAAGACATTTGGTGGTCAAGGACCTCAGAGCTCAGAGGTGTGAGCAGCAGGA 2005
QY 338 TCTCGGAAAGCTCAGTGTATCAGATGAGTGGCCATCAGCAGTTCACCACTTGAAGTGC 337
Db 2006 TCACAGAAAGCTGGAGATCAGACGACAGATTTGGGAAGCATCGAGTCTCATGCA 2065
QY 398 ATGCTAATGGTGTCCCGAGCCTCAGATCAGTCTGGTTTAAACACACCAAAATACAC 457
Db 2066 CGGCACTGGGAATCCCTCCACAGATCATGTGTTTAAAGATATGAGACCTTGTAG 2125
QY 458 AAGACCTTGGAAATATTTAGACAGCAGCAGCAGCAGCTGTTTATGAGAGCTGACAG 517
Db 2126 AAGACTCAGGCAATGATTAAGAGGATGGGAACCGGAACCTCAGTATCCGAGAGTAGGA 2185
QY 518 AAGAGGATGAGGTGTCTATCAGTCCAAAGCCACCAACAGAGAGGCTCTGGAAGAT 577
Db 2186 AGGAGGAGAGGCTCTACACTCCAGGATGCACTGCTGTGCAAAAGTGG 2245
QY 578 CAGCATCTCAGTCTTCAAGAACCTCGACACAGTCTAATCTGAGCTGATCATCTAA 637
Db 2246 AGGCAATTTTCAATATAGAGGTGCCAGAAAGACGAACTTGGAAATCATTTTCTAG 2305
QY 638 CATGCACTGTGGTGGGACTCTCTTCTGCTCTCTAATACCCTCTTATCCCAAAA 697
Db 2306 TAGGCAGGGGTATGGCATGTTCTTCTGGCTACTTCTGATCATCTTACCGAGCC 2365
QY 698 TGAAGAGTCTCTTCTGCTGAAATGAAGTCTACTACCTATCAATTAATATGAGCCAG 754
Db 2366 TTAAGCGGGCAATGAGAGGGAACGAGACAGAGGTACCTGTCCATGCTGATGATCCAG 2425
QY 755 ATGAGATTCCTTTGATGAGCAGTGTAGCGGCTCCCTTAATGATCCAGCAGTGGAGT 814
Db 2426 ATGAATCTCCCAATGGATGATGATGAGAGTGTGAGAGTGTGCTTAATGATGCAAAATG 2485
QY 815 TTGCGCGGAGAGACTTAAATGCGGCAATCACTTGGAGAGGGCTTTTGAAGATGG 874
Db 2486 TCCGAGAGCGGTGAGCTAGTAGAGCTCTTGGCTGTGGCTTTGGCAAGTGA 2545
QY 875 TTAAGCATCAGATTTGGCATTAAGAATCACTAGTCCGAGGACTGTGCTGTGAAA 934
Db 2546 TTAAGCAGAGATGCTTTGGAATTTGAAGACAGACGAACTTCCAGGACAGTAGCAGTCAAA 2605
QY 935 TCTGAAGAGGCGCCAGCGGACGAGTACAAAGCTCTGATGAGCTGATTAATCT 994
Db 2606 TGTGAAAGAGAGAGACACACAGAGTGAATGAGTCACTGCTGAGCTCACTGCTGAGTCA 2665
QY 995 TGACCCCATTTGCTGACCTGATCTGAGTGTGTTAACCCTGCGAGCTCCACCAAGCAAG 1054
Db 2666 TCATTCATTTGCTGACCTCAATGTTGTTCAACCTTCTAGGTGCTGTACCAAGCCAG 2725
QY 1055 GAGGCTCTGATGCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1114
Db 2726 GAGGCTCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2785
QY 1115 AGAGCAAAAGTGTATTTTCTCAACAGGATGAGCAGCTACACATGAGAGCTAAGA 1174
Db 2786 GAGCAGAGAGAGATGATTTGCTCTACAGACCAAGGGGACAGATCTGCTGAGGGA 2845
QY 1175 AAGAAAAATGAGGCGAGGCTGAGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1234
Db 2846 AAGACTAGTTGAGCAATCCCTG---TGGATCTGAAAGCGGCTTGGACAGCATCA 2902
QY 1235 GAGGCAAGCTTTGAGGCTCGGCTTTTACAGGAGATTAAGATGATGATGATGATGATGAT 1294
Db 2903 GTAGCAGAGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2962
QY 1295 AAGAGGAGATTTGAGGCTTTTACAGGAGGCGGCTTACAGGAGCTGATGATGATGATGAT 1354
Db 2963 AAGAGGAGCTTTGAGGCTTTTACAGGAGCTTACAGGAGCTTACAGGAGCTTACAGGAG 3022

QY 1355 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1414
Db 3023 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCACAGG 3082
QY 1415 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1474
Db 3083 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3142
QY 1475 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1534
Db 3143 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3202
QY 1535 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1594
Db 3203 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3262
QY 1595 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1654
Db 3263 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3322
QY 1655 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1714
Db 3323 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3382
QY 1715 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1774
Db 3383 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3442
QY 1775 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1834
Db 3443 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3502
QY 1835 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 1899
Db 3503 ACATTTTCAATGGCCAGAGCATGAGTCTTCTGCTTCCAGAAAGTCAATTCATCGG 3527

RESULT 2

US-09-483-539-1
; Sequence 1, Application US/09483539
; Patent No. 6359115
; GENERAL INFORMATION:
; APPLICANT: Kendall, Richard L.
; APPLICANT: Thomas, Kenneth A.
; APPLICANT: Mao, Xianzhi
; APPLICANT: Tobben, Andrew J.
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Merck & Co., Inc.
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/483.539
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19963PV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732/594-3905
; TELEFAX: 732/594-4720
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 4071 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-483-539-1

Query Match 34.0%; Score 713.8; DB 4; Length 4071;
Best Local Similarity 65.5%; Pred. No. 8.5e-212;
Matches 1077; Conservative 0; Mismatches 562; Indels 5; Gaps 2;

Qy 218 TCAATGATGTTCCCTGCAAGATTCAGGCACCTATCCCTCCAGAGCCAGGAATGTATACA 277
Db 1886 TTAAGATGATCTCTTGGAGCAACAGAGACTATGTGCTTCTCAAGACAGAGA 1945
Qy 278 CAGGGGAAGAAATCCTCCAGAAAGAAATTAACAATCAGAGATCAGGAAGACCAATACC 337
Db 1946 CCAAGAAAGACATTCCTGGTGGAGCAGCTCACAGCTCTAGAGCGTGTGGCAACCA 2005
Qy 338 TCCGTGGAACCTCAGTGAACAGACAGTGGCCATCAGAGTTCACCACTTTAGACTGTC 397
Db 2006 TCACGGAACCTGGAGATCAGAGCAAGATTTGGGAAAGCATCGAAGTCTCAIGCA 2065
Qy 398 ATGCTAATGTTGTCCTCCGAGCTCAGATCATTGTTTAAAGAACACCAAAATACAAC 457
Db 2066 CGGATCTGGAAATCCCTCCAGATCAATGTTTAAAGATTAAGACCCCTGTAG 2125
Qy 458 AAGAGCTGGAATTTATAGGACAGGAGACAGCTGTTTATTGAAGAGTCCAG 517
Db 2126 AAGACTCAGCATTTGATTGAAGATGGGAACCGGAACCTCACTATCCGAGAGTCAGCA 2185
Qy 518 AAGAGATGAGGTGCTATCAGTCGAAGCCACACACAGAGAGGCTCTGTGGAAAGT 577
Db 2186 AGGAGCAAGAGGCTCTACACCTCCAGAGCTCAGTGTCTTGGTGTGCAAAAGTGG 2245
Qy 578 CAGCATACCTCAGCTGTCAAGAGACCTCGGACAGTCTAATCTGAGCTGATCAGCTAA 637
Db 2246 AGGCAATTTCAATAGAGGTCGCCAGGAAGAGCACTTGGAAATCAATATCTAG 2305
Qy 638 CATGACCTGTGTGGCTGCGACTCTTCTGTGGCTCTTATACCTCTCTTATCCGAAA 597
Db 2306 TAGCAGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2365
Qy 698 TGAAGAGTCTCTTCTGATTAAGAGTCTGACTTACTATCAATTAATGAGCCAG 754
Db 2366 TTAAGCGGCGCAATGGAGGAGCTGAAGACAGGTTACCTGCTCACTGATGCTCAG 2425
Qy 755 ATGAGTCTCTTGTGATGAGCTGTGAGGCGCTCCCTTATGATCCAGCAAGTGGAGT 814
Db 2426 ATGAACTCCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2485
Qy 815 TTGCGCGGAGAGACTTAACCTGGGCAATCAGTTGGAGAGGGGCTTTTGGAAAGTGG 874
Db 2486 TCCCGAGAGAGCGGTGAAGCTAGGTAAGGCTCTTGGCGGTGGTGGCTTTGGCGCAAGTCA 2545
Qy 875 TTCAAGCATCAGATTTGGCAATTAAGAAATCACTTACCTGCGGAGCTGTGGCTGTGAAA 934
Db 2546 TTGAGCAGATGCTTTTGAATTTGACAGACAGCACTTTCAGAGACAGTACAGTCAAA 2605
Qy 935 TCGTGAAGAGGCGCGCGAGCGAGTACAAAGCTCTGAAGTCTGAGTGAAGTAAATCT 994
Db 2606 TGTGAAGAGAGGCAACACAGATGAGCATGAGCTCTGATGCTGATGCTGATGCTGATGCT 2665
Qy 995 TGACCCACATTTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1054
Db 2666 TCAATCATATTTGTCACCACTCAATTTGGTCAACCTTCTAGTGGCTGTACCAAGCAG 2725
Qy 1055 GAGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1114
Db 2726 GAGGCGCACTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2785
Qy 1115 AGAGCAACCTGACTTATTTTCTCAACAGGATGACAGCTACACATGAGGCTAAGA 1174

Db 2786 GGAGCAAGAGAAATGAATTTGTCCCTTACAAGCAACAAAGGGCCAGATTCCTCGTCAAGGA 2845
Qy 1175 AAGAAAAAATGAGCGAGGCTGGAACAAAGAAACCAAGAACTAGATAGGTCAACCA 1234
Db 2846 AAGACTACGTTGGAGCAATCCCTG---TGATCTGAAGCGGCTTGGACAGCATCACA 2902
Qy 1235 GCAGCAAGCTTTTCCGAGCTCCGGCTTTTCAAGAGATAAAGTCTCAGTGTGTTGAGG 1294
Db 2903 GTAGCAGAGCTCAGCCAGCTCTGGAATTTGTGGAGGAGAAGTCCCTCAGTGTGTGAAG 2962
Qy 1295 AAGAGCAGATTTCTCACGGTTTCTACAGGAGCCCATCACTATGAAGATCTGATTTCTT 1354
Db 2963 AAGAGAACTCTCTGAAGATCTGTATGAAGCTTCTGACCTTGGAGCATCTCATCTGTT 3022
Qy 1355 ACAGTTTTCAGGTGCCAGAGGAGTGGAGTCTCTGTTCTCCAGAAAGTGAATCATCGGG 1414
Db 3023 ACAGTTTTCAGGTGGCTAAGGCGATGGAGTTCTTGGCATCGGAAGTGTATCCACAGG 3082
Qy 1415 ACCTGGCAGGAGAAACATTTTATCTAGACACACGTTGTTGAAGATTTGTGATTTT 1474
Db 3083 ACCTGGCGGCAAGAAATATCTCTTATCGGAGAAAGAGTGGTTAAATCTGTGACTTTG 3142
Qy 1475 GCCTTCCCGGGATTTTATGAACCCCGATTTATGTGAAGAAAGAGATCTCGACTTC 1534
Db 3143 GCATGGCCGGGATTTTATGAAGATCCAGATTTATGTGAAGAAAGAGATCTCGCTCC 3202
Qy 1535 CTCTGAATGAGTGTCCGAACTATCTTTGACAAATCTACAGCAACAGAGAGGAGG 1594
Db 3203 CTCTGAATGAGTGTCCGCAAGAAATTTTTCACAGAGTGTACACAATCCAGAGTGAAG 3262
Qy 1595 TGTGCTTACGAGTATGCTGTGGGAAATCTTCTTGTAGTGGTCTCCATACCCAG 1654
Db 3263 TGTGCTTGTGTTTGTGTGGGAAATATTTCTTGTAGTGTCTTCTCAATCTCTG 3322
Qy 1655 GAGTACAAATGATGAGGACTTTTCCAGTCCCTGAGGAAAGGATGAGGATGAGAGTCT 1714
Db 3323 GGGTAAAGATGATGAAGAAATTTTGTAGGCGATTGAAGAAAGAGAACTAGAATCAGG 3382
Qy 1715 CTGAGTACTTCTCTGAATCTATCAGATCATGCTGAGCTGTGGCAGAGAGAGCCAA 1774
Db 3383 CTGATTAATACACCAAGAAATGACCAAGACCATGCTGAGCTGTGGCAGGAGCCCA 3442
Qy 1775 AAGAAAGCCAGATTTGCAAGCTTTGTGAAAGAACTAGGTGATTTGCTTCAAGCAATG 1834
Db 3443 GTCAGAGCCAGCTTTTCAAGTGTGGTGAACATTTGGAAATCTTGTGAAGCTAATG 3502
Qy 1835 TACAACAGGATGTTAAAGACTACAT 1859
Db 3503 CTCAGCAGGATGCAAGACTACAT 3527

RESULT 3
US-08-810-116-7
Sequence 7, Application US/08810116
Patent No. 5766860
GENERAL INFORMATION:
APPLICANT: Terman, Bruce I.
APPLICANT: Carion, Miguel E.
TITLE OF INVENTION: Identification of a No. 5766860el Human Growth
Factor Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS


```

: Sequence 7, Application US/07930548A
: Patent No. 5861301
: GENERAL INFORMATION:
: APPLICANT: Terman, Bruce I.
: APPLICANT: Carrion, Miguel E.
: TITLE OF INVENTION: Identification of a No. 5861301el Human Growth
: TITLE OF INVENTION: Factor Receptor
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: American Cyanamid Company
: STREET: One Cyanamid Plaza
: CITY: Wayne
: STATE: New Jersey
: COUNTRY: U.S.A.
: ZIP: 07470
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/930.548A
: FILING DATE: 23-NOV-1992
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Gordon, Alan M.
: REGISTRATION NUMBER: 30.637
: REFERENCE/DOCKET NUMBER: 31,298-01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-831-3244
: TELEFAX: 201-831-3305
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4236 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..4068
: US-07-930-548A-7

```

```

Query Match      33.8%; Score 710.5; DB 2; Length 4236;
Best Local Similarity 65.3%; Pred. No. 8.6e-211;
Matches 1075; Conservative 0; Mismatches 564; Indels 6; Gaps 2;

Y 218 TCATGAATGTTCCCTGCGAAGATTCAGGCACCTATGCTGCGAGCCAGGAAATGATACA 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1886 TTAAGATGCAATCTTGCGAGGACCAAGGAGACATATGCTGCTTGTCAAGCAGGAGA 1545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 278 CAGGGAGAGAAATCCCTCCAGAAGAAAGAAATTAACATCAGACATAGGAAGCACCATACC 337
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1946 CCAAGAAAGACATTCGGTGGTTCAGGCAGCTCAGAGTCCATAGAGGCTGTGCGCACCAGA 2005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 338 TCTGCGAAACCTCAGTATCAGACATGAGTGGCCATCAGCAGTTCACCACTTACAGCTGC 397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2006 TCACAGNAACCTTGGAGATCAGACGACAGATATGGGAAACCATCGAGTCTATGCA 2065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 398 ATGCTAATGTTGCCCGAGCTCAGATCAGTGGTTTAAACAAACCAACAAATACAAAC 457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2066 CGGCATCTGGGAATCCCTCCACAGATCATGTGGTTTAAAGATAATGAGACCCCTTGAG 2125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 458 AAGAGCTTGGANATATTTTAGACACAGGAAGCAGCAGCTGTTTATTAAGAGATCAGAG 517
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2126 AAGACTCAGGCATTTGATTAAGAGATGGGAACCGGAAACCTCAGTATCCGAGAGTGA 2185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 518 AAGAGATGAGGTGCTATCACTCAAGGCCACCAAGCAGAGGCTCTGTGGAAAGTT 577
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2186 AGGAGCAGGAGGCTCTACACTCCAGGATGCGAGTGTCTTGGCTGTGCAAAAGTGG 2245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 578 CAGCATACCTCTGTTCAAGGAACCTCGGACAAAGCTTAATCTGAGGTGATCACTCTAA 637
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 2246 AGGCATTTTTCATAATAGAGGTGCCAGAGAAAGACGAACTTGGAAATCATTTTCTAG 2305
Y 638 CATGCACCTGTGTGGCTGGACTCTCTCTGCTCTCTTAACTCTCTTATTCGAAAAA 697
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2306 TAGGCACAGCGGATGTGCAATGTTCTCTGCTCTCTTGTGTCATCATCTAGGACCG 2365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 698 TGAAGAAGTIC---TTCCTTCTGAATAAAGACTGACTACCTATCAATTAATTAAGACCCAG 754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2366 TTAAGCGGGCAATGGAGGGGAACCTGAAGACAGGCTACTTGTCCATCGTCATGATCCAG 2425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 755 ATGAAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCCCTTATGATGCCAGCAAGTGGAGT 814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2425 ATGAATCCCAATTTGATGAACATTTGTAACGACTGCTTATGATGCCAGCAATGGGAAT 2485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 815 TTGCGCGGAGAGACTTAACTTGGGCAATCACTTGAAGAGGGGCTTTTGGAAAGTGG 874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2486 TCCCAGAGACCGGCTGAACCTTAGTAAAGCTCTTGGCGGTGTGCTTGTGCCAGAGA 2545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 875 TTAAGCATCAGCAATTTGGCATTAAGAAATCACTAGTGCAGGACTGTGCTGTGAANA 934
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2546 TTAAGCAGATGCTTTTGAATTTGACAAGACAGCAACTTGCAGGACAGTAGCAGTCAAA 2605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 935 TCGTGAAGAGGGGGCCAGCCGACGAGTAAAGCTCTGTGACTGAGCTAAATCT 994
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2606 TGTGAAGAGAGGAGCACACACAGTGAAGCTCTCATGTCTGAACCTCAAGATCC 2665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 995 TGACCCACATTTGCCACCATCTGAACCTGTGTTAACTTGTGGAGGCTTGCACCAAGCAAG 1054
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2666 TCATTATATTTGTCACCATCTCAATGTGTTGTTCACTTCTAGTGCCTGTACCAAGCAG 2725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1055 GAGGCTCTGTGTTGATTTGAAATCTGCAATATGGAATCTCTCAACTACCTCA 1114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2726 GAGGCCACTCATGTTGTTGGAATTCGAAATTTGGAACCTTCCACTTACCTGA 2785
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1115 AGACCAAGCTGACTTATTTTCTCAAGAGTGCAGCACTACACATGAGCTTGA 1174
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2786 GGAGCAAGAGAAATGAAATTTGTCCTACAAAGCAAAAGGGGCGACGATTCGGTCAAGGA 2845
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1175 AAGAAAAATGGAGCCCGCTGGAAACAGCAAGAAACCAAGACTAGATAGCTCACC 1234
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2846 AAGACTAGTTGGAGCAATCCCTG---TGGATCTGAACGGGCTTGCACAGCATCACA 2902
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1235 CGAGGAAAGCTTTGGAGCTCGGCTTTTCAGGAAGATAAAGTCTGAGTGTGAGG 1294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2903 GTAGCGAGCTCAGCAGCTGTGATTTGTGAGGAGAGTCCCTCAGTGTATAGAAG 2962
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1295 AAGAGGAGTCTCTGAGGTTTCTACAAGAGGCCATCATATGGAAGATCTGATTTCT 1354
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2963 AAGAGGAGCTCTGAGATCTGTATAGGACTTCTGACCTTGGAGCATCTCATCTGT 3022
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1355 ACAGTTTTCAGTGGCCAGCAGCATGAGTTCTGTCTTCCAGAAAGTGCATTCATCGG 1414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3023 ACAGCTTCCAGTGGCTAAGGGCATGGAGTTCTTTGGCATCGGAAAGTATCCACAGG 3082
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1415 ACCTGGCAGCGAACAATCTTTTATCTGAGAACACAGTGGTGGAGATTTGTGATTTG 1474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3083 ACCTGGCGGCACCAATATCTCTTATCGGAGAACAGCTGGTTAAATCTGTGACTTG 3142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1475 SCCTTGGCCGGATATTTATAAGAACCCCGATATGTGAGAAAAGAGATCTCGACTTC 1534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3143 GCTTGGCCGGATATTTATAAGATCCAGATATGTGAGAAAAGGATGCTCGCTCC 3202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1535 CTCGAAATGAGTGGCTCCCGATCTCTTTGACAAATCTACAGCACCAGAGGAGG 1594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3203 CTTTGAATGGATGGCCCGCAGAAACAATTTTIGACAGAGTGTACACAATCCAGAGT 3262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1595 TGTGCTCTTACGAGTATGCTGTGGAAATCTTCTCTTAGTGGTCTCCATACCAG 1654
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3263 TCTGCTCTTGTGTTTGTCTGTGGAAATATTTTCTTAGTGGTCTCTCCATCTCG 3322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Y 1655 GAGTACAAATGAGTAGGACTTTTCCAGTCCCTGAGGGAAGCATGAGGATGAGCTC 1714
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3323 GGTAAAGATGATGAAGATTTTGTAGGCGATTTGAAAGAGGAACTAGAAATGAGGCC 3382
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY 1715 CTGAGTACTTCTGTAATCTATGATGATGCTGGAGCTGCTGGACAGAGACCA 1774
Db 3383 CTGATTTATATACACCAAAATGATACAGACCATCTGGACTCTGGACGGGAGCCA 3442
QY 1775 AAGAAAGCCAGATTTGAGAGACTTGGAAAGCTAGGTAATGCTTTCAGCAATG 1834
Db 3443 GTCAGAGACCCAGCTTTTCAGATGTTGGAGACATTTGGGAAATCTTTCAGCAATG 3502

QY 1835 TACAACAGGATGTAAGACTACAT 1859
Db 3503 CTCAGCAGGATGCAAGACTACAT 3527

RESULT 5

US-08-443-861-1

: Sequence 1, Application US/08443851

: Patent No. 5851999

: GENERAL INFORMATION:

: APPLICANT: Ullrich, Axel

: APPLICANT: Bisau, Werner

: APPLICANT: Millauer, Birgit

: APPLICANT: Gazit, Aviv

: APPLICANT: Levitzki, Alex

: TITLE OF INVENTION: Flk-1 Is A Receptor For Vascular

: TITLE OF INVENTION: Endothelial Growth Factor

: NUMBER OF SEQUENCES: 5

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Pennie & Edmonds

: STREET: 1155 Avenue of the Americas

: CITY: New York

: STATE: New York

: COUNTRY: U.S.A.

: ZIP: 10036-2711

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent Release #1.0, Version #1.25

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/443,861

: FILING DATE: 22-MAY-1995

: CLASSIFICATION: 514

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/193,829

: FILING DATE: 09-FEB-1994

: ATTORNEY/AGENT INFORMATION:

: NAME: Coruzzi, Laura A.

: REGISTRATION NUMBER: 30,742

: REFERENCE/DOCKET NUMBER: 7683-060

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (212)790-9090

: TELEFAX: (212)869-9741

: TELEX: 66141 PENNIE

: INFORMATION FOR SEQ ID NO: 1:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 5470 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: unknown

: TOPOLOGY: unknown

: MOLECULE TYPE: DNA

: FEATURE:

: NAME/KEY: CDS

: LOCATION: 286..4386

US-08-443-861-1

Query Match 33.8%; Score 710; DB 2; Length 5470;

Best Local Similarity 55.8%; Pred. No. 1.6e-210;

Matches 1081; Conservative 0; Mismatches 575; Indels 6; Gaps 2;

QY 201 CACTCTTAATCTTACCAATGATGTTTCCCTGCAAGATTCAGGCACCTATGCCCTGCAG 207
Db 2148 CATCTGATGTTGGCAATTCAGATGCCCTCTCTGAGGACAGGCGACTGTTGTTGCT 2077

QY 1338 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1397
DB 3285 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3344
QY 1398 AAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1457
DB 3345 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3404
QY 1458 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1517
DB 3405 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3464
QY 1518 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1577
DB 3465 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3524
QY 1578 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1637
DB 3525 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3584
QY 1638 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1697
DB 3585 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3644
QY 1698 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1757
DB 3645 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3704
QY 1758 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1817
DB 3705 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3764
QY 1818 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 1877
DB 3765 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGCGATGGAGTTCTGTGCTCCAG 3824

RESULT 6
US-08-193-829B-1
Sequence 1, Application US/08193829B
Patent No. 6177401
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Risaun, Werner
APPLICANT: Millauer, Birgit
APPLICANT: Gazit, Aviv
APPLICANT: Levitzki, Alex
TITLE OF INVENTION: FIX-1 Is A Receptor For Vascular
TITLE OF INVENTION: Endothelial Growth Factor
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 09-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzel, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-060
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-5090
TELEFAX: (212)869-5741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5470 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 286...4385
US-08-193-829B-1

Query Match 33.8%; Score 710; DB 4; Length 5470;
Best Local Similarity 65.0%; Pred. No. 1.6e-210;
Matches 1081; Conservative 0; Mismatches 575; Indels 6; Gaps 2;

QY 201 CACCTTAACTTACCATCATGATGTTTCCCTGCAAGATTCAGGCACCTATGCTGCTGAG 250
DB 2148 CATCTTGAATGGGCAATTCAGATGGCTCTCTGAGGACCAAGGCGACTATGTTTGCCT 2207
QY 261 AGCCAGGATGTATACACAGGGGAAGAAATCTCCAGAGAAAGAAATTAACAATCAGAGA 320
DB 2208 TGTCAAGATAAGAGACCAAGAAAGACATTGCTGGTCAACAGCTCATCATCTCTAGA 2267
QY 321 TCAGAGAGCCACCAATCTCTGCGAACTCAGTGATCAGACAGTGGCCATCAGCAGTTC 380
DB 2268 GCGATGCGCCCAATCATCCCGGAAATCTGGAGATCAGACAAACCAATTCGCGAGAC 2327
QY 381 CACCACTTATAGATGTCATGCTAAATGGTGTCCCGAGCTCTAGATCAGTCTGTTAAAA 440
DB 2328 CATTCAGAGTACCTGCGAGATCTGGAATCTTACCCACACATTCATGTTCAAGA 2387
QY 441 CAGCACAATAATACACAGAGCCCTGGAATTTTAGACAGGAGCAACAGCAGCTGTT 500
DB 2388 CAACGAGACCTTGGTAAAGATTCAGGCAATGTAAGAGATGGAAACCGGAACCTGAC 2447
QY 501 TATTCAGAGAGTACAGAGAGAGATGAGGTGTCTATCTGCAAGCCACCAACAGAA 550
DB 2448 TATTCAGAGAGTACAGAGAGAGATGAGGTGTCTATCTGCAAGCCACCAACAGAA 2507
QY 561 GGGCTCTGGAAGTTCAGATACCTCAGTCTTCAAGGAACTCGGACAACTGTAATCT 620
DB 2508 TGGCTGTGCAAGAGCGAGAGCTCTTCAATAGAGAGTGGCCAGGAAAGACCACTT 2567
QY 621 GAGCTGATCAGTCTAAGATGACAGCTGTGGCTGGAGCTCTCTTGGCTCTTATTAAC 680
DB 2568 GGAAGTCAATTAATCTGCGGACCTGAGTGGATGATGCTTCTGCTCTCTTCTTGT 2627
QY 681 CTCTCTTATCCGAAATAAGAAAGTCTCTCTGTAATAAGAACTGACTACTATC 737
DB 2628 CATTCGCTACGAGCTTACGCGGCAATGAGGGAAGTGAAGACAGAGCTACTTCTC 2687
QY 738 AATTATAAGGACCCAGATGAGTCTTGGATGAGCAGTGTGAGCGCTCTCTTATGA 797
DB 2688 TATTGATGATCCAGATGAATTCCTTGGATGAGCGCTGGAACCTTGGCTTATGA 2747
QY 798 TCCAGAGAGTGGGAGTTTCCCGGAGAGAGATTAACCTGSCAAATCAGTGAAGG 857
DB 2748 TCCAGAGAGTGGGAGTTTCCCGGAGAGAGATTAACCTGSCAAATCAGTGAAGG 2807
QY 858 GSCCTTTGAAAGTGGTTCAAGCATCAGCATTTGGCAATTAAGAAATCAGTCTGCTCG 917
DB 2808 TGCCTTCGCGCAGATGATTCAGGAGAGCTTTTGAATGACAGAGAGCTACTTCTCA 2867
QY 918 GACTGTGGCTGTGAATGCTGAAGAGGGGCGCCAGCGGAGTGAAGCTCTGAT 977
DB 2868 AACAGTGGCTGAAGTGTGAAGAGAGGAGCAACACAGAGAGCTACTGAGCGCTCAT 2927
QY 978 GACTGAGCTAATAATCTGACCCAGATTCGCCACCATCTCAACCTGTTACCTGCTGG 1037
DB 2928 GCTGAGACTCAAGATCTTCATCAGATTTGGTGGTGAACCTCTCTAGG 2987

QY	1038	AGCCTGCACCAAGCAGCAGGAGGCGCTCTGATGGTGTGATTTGAACTACTCAAAATATGGAAA	109
Db	2988	CGCCTGCACCAAGCGGAGGCGCTCTCATGGTGAATTTGGAATTTCTGCAAGTTTGGAAA	3047
QY	1098	TCTCTCCAATACCTCAAGAGGAAAGCTGTACTTATTTTCTCTCAACAGATCCAGCACT	1157
Db	3048	CTATCACTACTTACGGGGCAGAGAATGAAATTTGTTCCCTATAGAGCAAGGGGCG	3107
QY	1158	ACACATGGAGCCTAAGAAAGAAAATGGAGCCAGCCCTGGAAACAGGCAAGAAACCAG	1217
Db	3108	ACGCTTCCGCCCAAGGCAAGCACTACGTTTGG--GAGCTCTCGGTGATCTGAAAAGACG	3164
QY	1218	ACTAGATAGCTACACGACAGGAAAGCTTTTGGAGCTCCGSCCTTCAGCAGATAAAG	1277
Db	3165	CTGGACAGCATCACACAGCAGCAGCTCTGCCCTCAGGCTTTGTTGAGGAGAAATC	3224
QY	1278	TCTGAGTGTGTTGAGGAAGAGGAGATTTCTGACGGTTCTTACAAGGAGCCATCACTAT	1337
Db	3225	GCTCAGTGTGTAGGAAGAAAGAGCTTCTGAAGAACTGTAGAAGCACTTCAGCTT	3284
QY	1338	GGAGATCTGATTTCTTACAGTTTTCAAGTGGCCACAGSCATGGAGTTCTGTCTTCCAG	1397
Db	3285	GGAGCATCTCATCTGTACAGCTTCAGAGTGGCTAAGGCGATGGAGTTCTTGGCATCAAG	3344
QY	1398	AAAGTGCAATCTCGGACCTTGGCAGCGAGAAACATCTTTTATCTGAGAAACAAGTGCT	1457
Db	3345	GAAGTGTATCCACAGGACCTGGCAGCAGAAATCTCTCTATCGGAGAGAACTGCTGT	3404
QY	1458	GAAGATTGTGATTTGGCCTTGGCCGGGATATTTATAGAACCCTGATTTATGTGAGAAA	1517
Db	3405	TAAGATCTGTGACTTGGCTTGGCCGGGCAATTTATAAGACCCGGATTAITGTGAGAAA	3464
QY	1518	AGGAGATCTCGACTTCTCTGAAATGGATGGCTCCGGAATCTATCTTTCAGAAATCTA	1577
Db	3465	AGGAGATGCCAGCTCTCTTGAAGTGGATGCCCGGAACCAATTTTGCACAGAGTATA	3524
QY	1578	CAGCACCAAGAGCAGCTGTGGTCTTACGGAGTATTCGTGTGGGAAATCTTCTCTTAGG	1637
Db	3525	CACAATTCAGAGCGATGTGGTCTTTCGGTGTGCTCTGGGAAATTTTCTCTTAGG	3584
QY	1638	TGGTCTCCATACCCAGGAGTACAATCGATGAGGACITTTTCAGTCCCTGAGGAGG	1697
Db	3585	TGCTTCCCATACCTCGGGGTCAAGATTGTAGAGAAATTTGTAGGAGATTGAAAGAGG	3644
QY	1698	CATCAGGATCAGAGCTCTCTGAGTACTCTACTCTCTGAAATCTATCAGATCATGCTGGACTG	1757
Db	3645	AACTAGAATCGGGGTCTCTCACTACACTACCCAGAAATGTACCAGACCACTCGGACTG	3704
QY	1758	CTGGCAGAGAGCCCAAGAAAGGCCAAGATTGAGAGCTTGTGGAANAACATAGTGA	1817
Db	3705	CTGGCATGAGGACTCCCAACAGAGACCTCTGTTTCAGAGTTGGTGGAGCAITTTGGGAA	3764
QY	1818	TTTCTTCAAGCAATTTACAAGCAGTGGTAAAGACTACAT	1859
Db	3765	CTCTCTCAAGCAATTCGCGAGGAGTGGCAAGACTATAI	3805

RESULT 7

US-07-813-593-3

: Sequence 3, Application US/07813593

: Patent No. 5185438

: GENERAL INFORMATION:

: APPLICANT: Lemischka, Ihor R.

: TITLE OF INVENTION: TCTIPOTENT HEMATOPOIETIC STEM CELL

: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

: NUMBER OF SEQUENCES: 4

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: IMCLONE SYSTEMS INCORPORATED

: STREET: 180 VARICK STREET

: CITY: NEW YORK

: STATE: NEW YORK

: COUNTRY: U.S.A.

: ZIP: 10014

Db 2490 GAGAGTCAATATCTCGTGGGACCTGAGTGTGATGCTATGCTTCTTGGCTCCCTTCTGT 2549
QY 681 CTTCTTATCCGAAATGAAAGGTC---TTCTTCTGAAATAAGAGACTGACTACCTATC 737
Db 2550 CATCTCGTACGAGCGGTTAAGCGGGCAATGAAGGGGAACAGACAGAGCTACCTTGC 2609
QY 738 AATTAATAGGACGAGATGAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCCCTTAAGA 797
Db 2610 TATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2669
QY 798 TGCCAGCAAGTGGGAGTTTCCCGGAGAGACTTAACTGGGCAATACACTTGGAGAGG 857
Db 2670 TGCCAGCAAGTGGGAAATCCCGAGGAGCGGCTGAACTAGGAAACCTTGGCCCGG 2729
QY 858 GCTTTTGGAAAGTGGTTCAGCAATCAGCAATTTGGCAATTAACAAATCACTTCTCCG 917
Db 2730 TGCTTGGGCAAGTGTGAGGAGAGGCTTTTGAATTAACAGAGAGGAGCTTGA 2789
QY 918 GACTGTGCTGTGAATCTGAGAGGGGCGGCGGCGGAGTACAGCTTCAAT 977
Db 2790 AACAGTAGCGTCAAGATGTTGAGAGAGGAGCAACACAGCAGCAATCGAGGCTTCA 2849
QY 978 GACTGAGTAAATTTGACCCACATTTGGCCACATTCGAGCTGGTAACTTCTCTGG 1037
Db 2850 GTCTGACTCAAGATCTCAATGCTGACCACTCAATGCTGAGCTTCTCTGG 2909
QY 1038 ACCCTGACCAAGAGAGGCGCTTCTGATGATGATGATGATGATGATGATGATGATG 1097
Db 2910 CCGCTGACCAAGCGGAGGCGCTTCTGATGATGATGATGATGATGATGATGATG 2969
QY 1098 TCTCTCACTACTCTCAAGGCAAGCTGACTTATTTTCTCAAGAGAGTCACT 1157
Db 2970 CTTATCACTTACTTACGGGCAAGAGAAATGAAATTTCTCTATAGCAAGGCGC 3029
QY 1158 ACAGATGAGGCTTAAAGAGAAATTTGAGCCAGGCTTGAACAGCAAGCAAGCAAG 1217
Db 3030 AGCTTTCGCGGAGGAGGAGTGTGGG---GAGCTTCGCTGGATCTGAAGAGG 3086
QY 1218 ACTAGATAGGCTACAGCAGGAGGAGCTTTCGAGCTTCGCTTTCAGGAGATAAG 1277
Db 3087 CTGGACAGCAATCAGCAGCAGGAGCTTTCGAGCTTCGCTTTCAGGAGATAAG 3146
QY 1278 TCTGAGTGTGTTGAGAGAGAGGAGTCTGAGGTTTCTCAAGGAGGCTTCACTAT 1337
Db 3147 GCTCAGTGTGAGAGAGAGAGTCTTCTGAGAGTGTACAGGAGTCTTCTCAGCT 3206
QY 1338 GAGAGTGTGATTTTACAGTTCAGTGGCCAGAGGAGTGTGAGTCTTCTCTCAG 1397
Db 3207 GAGCATCTCATCTGTTCAGCTTCCAGTGGCTTAAGGCGATGAGTCTTCTGAGTCAAG 3266
QY 1398 AAGTGTGATTCATGGGAGCTGGCAGGAGAAATCTTTTATCTGAGAAACAGCTGT 1457
Db 3267 GAGTGTATCCAGAGGAGCTGGAGAGCAATCTTCTCTTATCGAGAGAGTGTGT 3326
QY 1458 GAAGATTTGATTTTGGCTTTCGCGGATATTTATAGAGCCGCTTATGAGAGAA 1517
Db 3327 TAAGATCTGTGACTTGGCTTGGCGGAGCAATTTAAGAGCCGCTTATGAGAGAA 3386
QY 1518 AGAGATCTGATCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1577
Db 3387 AGAGATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3446
QY 1578 GAGCAGCAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1637
Db 3447 CACATTCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3506
QY 1638 TGGGTCTTCAATCCAGAGAGTAAATGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 1697
Db 3507 TGGCTTCCCAATCCAGAGAGTAAATGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 3566
QY 1698 CATGAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1757

Db 3567 AACTAGAAATGGGCTCTGACTACACTACCCAGAAATGTACCAGACCATCTGGACTG 3626
QY 1758 TTGGCAGAGAGCCCAAGAAAGAGCCCAAGATTTGCAAGAACTTGTGAAAACTAGGTGA 1817
Db 3627 CTGSCATGAGGACCCCAACACAGAGACCTCTCTTTTTCAGATTTGGTGGACCATTTGGGAAA 3686
QY 1818 TTGCTTCAGCAAAATGTACACAGAGTGTGAAGACTATAT 1859
Db 3687 CTTCTCAAGCAAAATGCGCAGCAGSATGCGCAAGACTATAT 3728
RESULT 8
US-07-977-451-5
: Sequence 5, Application US/07977451
: Patent No. 5270458
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: TOTOPOENT HEMATOPOIETIC STEM CELL
: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ImClone Systems Incorporated
: STREET: 180 Varick Street
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10014
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/977,451
: FILING DATE: 19921119
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US UNASSIGNED
: FILING DATE: 12-NOV-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/906,397
: FILING DATE: 26-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCI/US92/05401
: FILING DATE: 26-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: TW 81102961
: FILING DATE: 15-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCI/US92/02750
: FILING DATE: 02-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/813,593
: FILING DATE: 24-DEC-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/793,065
: FILING DATE: 15-NOV-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/728,913
: FILING DATE: 28-JUN-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/679,666
: FILING DATE: 02-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Feit, Irving N.
: REGISTRATION NUMBER: 28,601
: REFERENCE/DOCKET NUMBER: LEM-3-7P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5406 base pairs

STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: FICOPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,507
FILING DATE: 19920917
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/579,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
NAME/KEY: mat_peptide
LOCATION: 208..4308
US-07-946-507-3

Query Match 33.7%; Score 708.4; DB 1: Length 5406;
Best Local Similarity 65.0%; Pred. No. 4.9e-210;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2:
QY 201 CACTCTTAATCTTACCATCATGAATGTTTCCTGCAAGATTTCAGGCACTATGCGTCAG 250
DB 2070 CACTCTTAATCTTACCATCATGAATGTTTCCTGCAAGATTTCAGGCACTATGCGTCAG 250
QY 261 ACCCAGGAGATGATACACAGGGGAGAAATCCTCCAGAAAGAAATTAACATCAGAGA 320
DB 2130 TCTCAAGATAG 2189
QY 321 TCAGGAG 380
DB 2190 GCGGATGCGGAG 2249
QY 381 CACCAG 440
DB 2250 CATTGAAGTGAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2309
QY 441 CAACACAG 500
DB 2310 CACAG 2369
QY 501 TATTGAAG 550

DB 2370 TATCCGAG 2420
QY 561 GCGCTCTGTCGAAAGTTCAGCATACCTCACTGTTCAGGAACCTCGGACAGAGTCAATCT 520
DB 2430 TGGCTGTGAG 2480
QY 621 GAGCTGATCACTCTTAACATGACCTGTGGGTGGGAGTCTCTTCTGGCTCTCTATTAAC 680
DB 2490 GGAAGTCAITATCTCGTCGACCTGAGTGAATGGCATGTTCTTCTGGCTCTCTTGT 2549
QY 681 CTTCTTATCGGAAATGAAAGGTC---TTCTCTGAAATAAAGACTGACTACCTATC 737
DB 2550 CATCTCGTACGAGCGGTTAGCGGGCAATGAGGGGAGACTGAGAGAGAGAGAGAG 2609
QY 738 AATTATATGACCCAGATGAAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCCCTTATGA 797
DB 2610 TATTGTCAATGATCCAGATGAATGCGCTTGGATGAGCGCTGTGAACGCTTGCCTTATGA 2659
QY 798 TGGCAGCAATGGGAGTTCCTCGGGGAGAGACTTAACCTGGCAATCACTTGAAGAGG 857
DB 2670 TGGCAGCAATGGGAGTTCCTCGGGGAGAGACTTAACCTGGCAATCACTTGAAGAGG 2729
QY 858 GCGTTCGAAAGTGTTCAGAGCATCAGCATTTTGGCAATTAAGAAATCAGCTAGTCCG 917
DB 2730 TGGCTTCGCGCAAGTGAATGAGCAGAGCGCTTTGGATGAGCAGCAGCAGTTCGAA 2780
QY 918 GACTGTGCTGTGAAATGCTGAAAGAGGGGCCACGCGGACGAGTACAAAGCTCTGAT 977
DB 2790 AACAGTACCGGTCAAGATGTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2840
QY 978 GACTGAGCTAATAATCTTGACCCACATTTGCCACCATCTGAGCGTGTACCTGCTGG 1037
DB 2950 GTCTGAATCAAGATCTCATCCATTTGGTCACTCTCAATGTTGGTGAACCTCTTAGG 2909
QY 1038 AGCTTCGACCAAGCAGAGAGGCGCTCTGATGTTGTTGAATTAATCTGCAATATGAAA 1097
DB 2910 CGCTTCGACCAAGCAGAGAGGCGCTCTCATGTTGTTGAATTTCTGAAAGTTGAAA 2969
QY 1098 TCTCTCACTACTCTAAGAGCAAGCTGACTTATTTTCTCAACAGGATGAGCAGCT 1157
DB 2970 CCTATCAACTTACTTACGGGGCAGAGAAATGAATTTGTTCCCTAATAGCAAGGGGC 3020
QY 1158 ACAGATGAGCTTAAGAAAGAAATGAGCAGGCGCTGGAACAGAGAGAGAGAGAGAG 1217
DB 3030 AGCTTCGCGCAGGCAAGAGTACGTTGGG---GAGCTCTCGTGGATCTGAAAGAGC 3086
QY 1218 ACTAGATAGCTCACCAGCAGCAGAGCTTTTCGAGCTCCGGCTTCAGGAGATNAAG 1277
DB 3087 CTTCGACAGATCACCAGCAGCAGAGCTCTGCCAGCTCAGGCTTGTGTGAGAGAAATC 3146
QY 1278 TCTGAGTGAATGAGAGAGAGAGAGATTTCTGAGGTTTCTACAGGAGGCCATCACTAT 1337
DB 3147 GCTCAGTGAATGAGAGAGAGAGAGAGCTTCTGAAGAGTGTGACAGGAGCTTCTGAGCT 3206
QY 1338 GGAAGATCTGAATTTTACAGTTTTCAAGTGGCAGAGAGAGATGAGTCTCTCTCCAG 1397
DB 3207 GAGGATCTCTGATCTGATGAGCTTCAAGTGGCTGAGGAGTGGTCTTGGCATCAAG 3266
QY 1398 AAGTGGATCTGAGGAGCTGCGAGCGGAGAGAGATTTTATCTGAGAGAGAGAGTGGT 1457
DB 3267 GAAGTGAATCTCAGAGGAGCTGCGAGCAGAGAGAGATTTCTCTTACGAGAGAGATGGT 3326
QY 1458 GAAGTGAATCTGAGTTCGCGGAGATTTTATGAAGAGAGAGAGATTTATGAGAGAG 1517
DB 3327 TATGATCTGAGTTCGCGGAGATTTTATGAAGAGAGAGATTTATGAAGAGAGAGAT 3386
QY 1518 AGAGATGAGTCTGCTGAAATGAGTGGCTCCGGAATCTATCTTTGAGAGAGAGATCT 1577
DB 3387 AGAGATGAGTCTGCTGAAATGAGTGGCTCCGGAATCTATCTTTGAGAGAGATCT 3446
QY 1578 CAGCAG 1637

Db 3447 CACAAATCAGAGCGAATGTTGGCTTTCCTGGAATAATTTCTTAGG 3506
QY 1638 TGGGTCTCCATACCCAGAGTACAAATGGATGAGGACTTTCCAGTCCCTGAGGGAAGG 1697
Db 3507 TGGCTCCCATACCTCGGGTCAAGATTGATGAAGAAATTTTGTAGGAGATTGAAGAAGG 3566
QY 1698 CATGAGATGAGAGCTGCTGAGTACTTCTGTAATCTATGAGATCATGCTGGAGCTG 1757
Db 3567 AACTAGAAATGCGGCTCTCTGACTACATACCCAGAAATGTACAGACCAATGCTGGAGCTG 3626
QY 1758 CTGGCAGAGAGCTCCAAAGAAAGCCAGAAATTTGAGAACTTTGGAAAACTAGTGCA 1817
Db 3627 CTGGCATGAGGACCCCAACACAGAGACCTCTGTTTTCAGAGTTGGGAGCAATTTGGGAA 3686
QY 1818 TTGCTTCAAGCAATGTACAAAGGATGTTGTAAGACTACAT 1859
Db 3687 CCTCTGCAAGCAATGCGCAGAGGATGCAAGACTATAT 3728

RESULT 10

Sequence 5, Application US/09252517
Patent No. 5548065
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: IDENTIFICATION OF HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Imclone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,517
FILING DATE: 31-OCT-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 25-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCI/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IW 81102961
FILING DATE: 13-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCI/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.

REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
US-08-252-517-5

Query Match 33.7%; Score 708.4; DB 1; Length 5406;
Best Local Similarity 65.0%; Pred. No. 4.9e-210;
Matches 1080; Conservative 0; Mismatches 576; Indels 2;

QY 201 CACTTAACTTACCATCATGTAATGTTCCCTGCGAGATTTCGACCTATGCTTGGCAG 260
DB 2070 CATCTTGAATGTCGATTTTCAAGATGCTCTCTGAGGAGGAGGAGGAGGAGGAGG 2129
QY 251 AGCCAGCAATGTATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 320
DB 2130 TGTCAAGATAGAGAGCAAGAGAGAGATTCCTTGGTCAAGAGCTATCATCTTGA 2189
QY 321 TCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 380
DB 2190 GCGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2249
QY 381 CAGCAGCTTATGAGCTGTATGCTAATGTTGCTCCCGAGGCTCAGATCATCTTGGTTTAA 440
DB 2250 CATTTGAAGTGAATTCCTCCAGCACTGGAATCTTACCTCCCAACATATCATGTTTCA 2309
QY 441 CAACCAACAAATACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 500
DB 2310 CAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2369
QY 501 TATTTGAAGAGTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 560
DB 2370 TATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2429
QY 551 GGGCTCTGTGGAAAGTTTACGATACCTTCTTCAAGAGAGGAGGAGGAGGAGGAGGAGG 620
DB 2430 TGGCTGTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2489
QY 521 GGAGCTGATCATCTTAAACATGCACTGTGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGG 680
DB 2490 GGAAGTATATCTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2549
QY 591 CCTCTTATCCGAAATCAAAAGGTC---TCTTCTGAAATTAAGAGCTGACTACCTATC 737
DB 2550 CATTCGTACGAGCAGGTTAAGCGGCGCAATGAAGGGAGGAGGAGGAGGAGGAGGAGGAG 2609
QY 738 AATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 797
DB 2610 TATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2669
QY 798 TGCAGCAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857
DB 2670 TGCAGCAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2729

QY 858 GCGTTTGGAAAAGTGGTTCAGCAATCAGCAATTTGGCAATTAAGAAATCAGCTACGTCGG 917
Db 2730 TGCTTTGGGCAAGATGATTCAGGAGAGCGCTTTTGGAAATGACAGAGCACTTGGAA 2789
QY 918 GACTGGCTGTGAAATGCTCAGAGAGGGGCCAGCGCCAGGACGAGTACAGCTTCAT 977
Db 2790 ACAGTAGCGCTCAGAGTGTGAAGAGAGGAGCAACACAGGAGCATCGAGCCCTCAT 2849
QY 978 GACTGAGCTAAATATTTGACCCACCAATTTGGCCACCATCTGAAAGTGGTAACTGCTGG 1037
Db 2850 GTCTGAATCAAGATCTCATCCATCATGATGTCACCACTCAATGTTGGTCACTCTAG 2909
QY 1038 AGCTTGACAGCAAGCAAGGAGGCGCTCTCATGCTGATTTGTAATGCTGCAAAATGAA 1097
Db 2910 CGCTTGCAAGCAAGCGGAGGCGCTCTCATGCTGATTTGTAATGCTGCAAAATGAA 2969
QY 1098 TCTCCCACTACCTCAGAGCAAGAGTGTGACTTATTTTTCACAAAGGATGAGCAGC 1157
Db 2970 CCAATCACTTACTTACGGGGCAAGAGAAATGAAATTTTCCCTATAGAGCAAGGCGC 3029
QY 1158 ACAGTAGAGCTCAGCAAGCAAGCAAGCTTTCCGAGCTCGGCTTTTTCAGCAAGATTAAG 1277
Db 3030 AGCTTTCCGAGGCAAGCAAGTACCTTTGGG---CAGCTCTCGTGGATCTCAAAAGAGC 3096
QY 1218 ACTAGATAGCTCAGCAAGCAAGCAAGCTTTCCGAGCTCGGCTTTTTCAGCAAGATTAAG 1277
Db 3087 CTTCAGAGATCAGCAAGCAAGCAAGCTTTCCGAGCTCGGCTTTTTCAGCAAGATTAAG 1346
QY 1278 TCTGAGTATGCTGAGCAAGCAAGCAAGCTTTGAGGCTTTTTCAGCAAGCAAGCTTAC 1337
Db 3147 GCTCAGTATGCTGAGCAAGCAAGCAAGCTTTGAGGCTTTTTCAGCAAGCAAGCTTAC 3206
QY 1338 GGAAGATCTGATTTTACAGTATTTTCAAGTGGCCAGGAGCAAGTATTTTTCAGTATTC 1397
Db 3207 GGAGCACTCATCTCTTACAGTATTTTCAAGTGGCCAGGAGCAAGTATTTTTCAGTATTC 3256
QY 1398 AAGTGCATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCTT 1457
Db 3267 GAAGTATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCTT 3326
QY 1458 GAAGTATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCTT 1517
Db 3327 TAAGATCTGATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAG 3386
QY 1518 AGGATATCTGATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAG 1577
b 3387 AGGATATCTGATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAG 3446
QY 1578 CAGCAAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTT 1637
Db 3447 CACAATCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTT 3506
QY 1638 TGGGTCTCATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCAAG 1697
Db 3507 TGGGTCTCATCTCAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCAAG 3566
QY 1698 CATGAGGATGAGCTCTCAGTATCTTACCTCAGCAAGCAAGCTTTTTCAGCAAGCAAGCT 1757
Db 3567 AACTAGATGCGGCTCTCAGTATCTTACCTCAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTT 3626
QY 1758 CTGCAAGCAAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTT 1817
Db 3627 CTGCAAGCAAGCAAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTT 3686
QY 1818 TTGCTTCAAGCAAGTATCAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAG 1859
Db 3687 CTCTCTGCAAGCAAGTATCAGCAAGCAAGCTTTTTCAGCAAGCAAGCTTTTTCAGCAAG 3728
RESULT 11
US-07-906-397A-5
; Sequence 5, Application US/07906397A

Patent No. 5521090
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOPIOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCLONE SYSTEMS INCORPORATED
STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397A
FILING DATE: 19920626
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: IEM-3-PPPPPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANIL-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 208..4308
US-07-906-397A-5

Query Match 33.7%; Score 708.4; DB 1; Length 5406;
Best Local Similarity 65.0%; Pred. No. 4,9e-210;
Matches 1090; Conservative 0; Mismatches 576; Indels 2;

QY 201 CACTCTTATCTTACCATCATGATGTTTCCCTGCAAGATTCAGGCACTATGCTCTGAG 263
Db 2070 CATCTTGTATTTGGCATTTTCAAGATGCCCTCTCTCGAGGACCAAGGGGACTATGTTTCTC 2129
QY 261 AGCAGGAATGATATACAGAGGGGAAGAAATCTCCAGAGAAAGAAATTAACATCAGAGA 320
Db 2130 TGTCTAGATAGAGACCAAGAAAGACATTCCTGCTCAACAGCTCATCATCTCAGA 2189
QY 321 TCAGGAAGCAGCATACCTCTCTCTGCAAGCACTGATTCACACAGTGGCCATCAGGATTC 390


```

1 PRIOR APPLICATION DATA:
2   APPLICATION NUMBER: US 07/793,065
3   FILING DATE: 15-NOV-1991
4 PRIOR APPLICATION DATA:
5   APPLICATION NUMBER: US 07/728,913
6   FILING DATE: 28-JUN-1991
7 PRIOR APPLICATION DATA:
8   APPLICATION NUMBER: US 07/579,666
9   FILING DATE: 02-APR-1991
10 ATTORNEY/AGENT INFORMATION:
11   NAME: Feil, Irving N.
12   REGISTRATION NUMBER: 28,601
13 REFERENCE/DOCKET NUMBER: LEM-3-7P
14 TELECOMMUNICATION INFORMATION:
15   TELEPHONE: 212-645-1405
16   TELEFAX: 212-645-2054
17 INFORMATION FOR SEQ ID NO: 5:
18   SEQUENCE CHARACTERISTICS:
19     LENGTH: 5406 base pairs
20     TYPE: nucleic acid
21     STRANDEDNESS: double
22     TOPOLOGY: linear
23     MOLECULE TYPE: cDNA
24     HYPOTHETICAL: NO
25     ANTI-SENSE: NO
26     FRAGMENT TYPE: N-terminal
27     FEATURE:
28       NAME/KEY: CDS
29       LOCATION: 208..4311
30       FEATURE:
31       NAME/KEY: mat_peptide
32       LOCATION: 265..4308
33       FEATURE:
34       NAME/KEY: sig_peptide
35       LOCATION: 208..264
36 PS-08-601-891-5

```

Qy	561	CTCTCTATCCGAAATAAGAAAGTCT--TTCTCTGGAATAAGACTGACTACTACTATC	737
Qy	562	TTCTCTGGAATAAGACTGACTACTACTACTATC	738
Db	2550	CAITCTCGTACGGACCGTITAGCGGSCCAATAGAGGGGSACTGAAGAGGCTACTGTGTC	2609
Qy	738	AAITATAATGGACCCAGATGAAGTCTCTTTGGATGAGCAGTGTGAGCGGCTCCCTTATGA	797
Db	2610	TATGTTCATGGATCCAGATGAATGGCCCTTGGATGAGCGCTGTGAAGCCTTGCCTTATGA	2669
Qy	798	IGCCAGCAGTGGGAGTITTCGCGGGAGAGACITTAACCTGGGCAAACTACCTTGGAGAGG	857
Db	2570	IGCCAGCAGTGGGAAITCCCAAGGAGCCGGCTGAACCTAGGAAAGACCTCTTGGCCGCGG	2729
Qy	858	GGCTTTTGGAAAAGTGGTTCACGATCAGCACTTTGGCATTAAGAAATACCTACGTGCGG	917
Db	2730	IGCCTTCGGCCAGTGTATGAGCGACAGCCTTTGGAAATGACAGACCGGCTTGCNA	2789
Qy	918	GACTGTGGCTGTGAANAATGCTGAAGAGGGGGCCAGCGCCAGCGAGTACAAAGCTCTGAT	977
Db	2790	AACAGTAGCCGTCAAGATGTTGAAGAAGAGCAACACACAGCGAGCATGAGCCCTCAT	2849
Qy	978	GACTGAGCTAAAATCTTGACCACATTTGACCACATCTGAACGTGGTAACTCTCTGGG	1037
Db	2850	GTCTGAACCTCAGATCTCTATCCCATTTGTCACCATCTCATGTGGTGAACCTCTTAGG	2909
Qy	1038	AGCCTGCACCAAGCAAGGAGGGGCTCTGATGGTGATTTGGTAATCTGCAAAATGGA	1097
Db	2910	CGCTGCACCAAGCGGAGGGGCTCTCATGTGATTTGGAAATCTCGAAGTTTGGAA	2969
Qy	1098	TCTCTCCACTACTCAGAGCAAAAGCTGACTTATTTTCTCANCAGATGCAGCACT	1157
Db	2970	CTATCAACTTACTTACGGGCAAGAGAAATGAATTTGTTCCCTATAAGACAAAGGGG	3029
Qy	1158	ACACATGGAGCCTTAAGAAAGAAAATGAGCGAGGCTGGAAAGCAAGCAAGAACCAAG	1217
Db	3030	AGGCTTCCGCCAGGCGAAGNACTACGTTGGG--GAGCTCTCGGTGATCTGAAGAGG	3086
Qy	1218	ACTAGATAGCGTACCAGAGCGAAAGCTTTTGGAGCTCCGGCTTCGAGAGATAAAG	1277
Db	3087	CTTGGACAGCATCACACAGCGACAGAGCTCTGCCAGCTCAGGCTTTTGGAGAGAAATC	3146
Qy	1278	TCGAGTGATTTGAGGAAGAGAGATTTGAGGGTTCTACAGGAGGCCATCAGTAT	1337
Db	3147	GCTCAGTGATGTAGAGGAGAGAGCACTTCTGAAGACTGTACAGAGACTTCTGACCT	3206
Qy	1338	GGAGATCTGATTTCTACAGTTCAGAGTGGCCAGAGGCTGAGTTCCTGTCTCCAG	1397
Db	3207	GGAGCATCTCATCTGTACAGCTTCCAGTGGCATAGGCGATGGAGTCTTGGCATCAAG	3266
Qy	1398	AAAGTGATTCACGGAGCTTGCACGAGAGAACTTCTTTATCTGAGAACACGTGT	1457
Db	3267	GAAGTGATCCACAGGAGCTTGCACAGCAAGAACTCTCTATCGAGAGAAATGTGT	3326
Qy	1458	GAAGATTTGTGATTTGGCCCTTGGCCGGATTTTATAGAACCCCGGATTTATGTGANA	1517
Db	3327	TAAGATCTGTGATTTGGCCCTTGGCCGGGACATTTAAGAACCCGGATTTATGCAAA	3386
Qy	1518	AGGAGATACCTGACATCTCTGAAATGGATGGCTCCGGAATCTATCTTTGCAAAATGA	1577
Db	3387	AGGAGATGCCGACTCCCTTTGAAGTGAATGGCCCGGAAACCACTTTTGACAGAGTAT	3446
Qy	1578	CAGCACCAAGAGCGAGTGTGCTTTACGGAGTATTTGCTGTGGGAAATCTTCTCTTAGG	1637
Db	3447	CACATTCAGCGGATGTGGTCTTCGGTGTGTGCTCTGGAAATTTTCCCTTAGG	3506
Qy	1638	TGGTCTCCATACCCAGGAGTACAAATGGATGAGAGCTTTTTCAGTCCGCTGAGGAGG	1697
Db	3507	IGCCTTCCCAATCCCTGGGGTCAAGATTCATGAAGAAATTTTGTAGAGATTTGAAGAGG	3566
Qy	1698	CATGAGATGAGAGCTCTCTGAGTACTACTCTGTAATCTATCAGATCATGCTGCAGTG	1757
Db	3567	AACTAGAAATGGGGCTCTCTGACTACTACTACCCCAAGAAATGCTAGAGCACTCTGAGTG	3626
Qy	1758	CTGGCACAGACCCCAAGAAAGAGGCAAGATTTTGCAGAACTTTGTGAAAACCTAGTGCA	1817

Db 3627 CTGGCAAGGACCCACACAGACACCTCGTTTCAGAGTGGTGAGCAATGGGAAA 3686
Qy 1818 TTGCTTCAAGCAATGTACACAGGATGTAAAGACTACAT 1859
Db 3687 CCTCTGCAAGCAATGTACACAGGATGTAAAGACTACAT 3728

RESULT 13

US-09-021-324-5

Sequence 5, Application US/09021324

Patent No. 5912133

GENERAL INFORMATION:

APPLICANT: Lemischka, Ibor R.

TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/021,324

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/977,451

FILING DATE: 1992-11-19

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/906,397

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/813,593

FILING DATE: 24-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/793,065

FILING DATE: 15-NOV-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/728,913

FILING DATE: 29-JUN-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Felt, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 5406 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

Qy 201 CACTCTTATCTTACCATCATGATGTTTCCCTGCAAGATTTCAGGCACCTATGCTGTGAG 260
Db 2070 CATCTTGATTTGGCAATTCAGAAATGCCCTCTCTGCAAGCAAGGCGACTATGTTGCTC 2129
Qy 261 AGCCAGGAAATGTATACACAGGGGGAAGAAATCCCTCCAGAGAGAAAGAAATTAACAATCAGAGA 320
Db 2130 TGCTCAAGATAGAGAGACCAAGAAAGACATTCCTTGTCACACAGCTCATCATCTAGA 2189
Qy 321 TCAGGAGACCATATACCTCTCTGCAAGACCTAGTGATCACACAGTGGCCATCAGAGTTC 380
Db 2190 GCGATGCGACCCCATGATCAGCGGAAATCTGGAGAAATCTGGAGAAATCAGACAAACCATTTGCGGAGAC 2249
Qy 381 CACCACCTTATGACTGTGATGCTAATGGTGTCCCGAGCTCAGATCACTTGTTTAAAAA 440
Db 2250 CATGAGTCACTTCCAGCATCTGGAAATCTACCCACACATTAATGGTTTCAAGA 2309
Qy 441 CAACCAAAATACAAAGAGCCCTGGAAATTTTATAGCAGCAGGAGCAGCAGCTGTT 500
Db 2310 CAAGAGACCTGTTAGAGATTCAGGCATTTGCTGAGAGATGGGAACCGGAACCTGAC 2369
Qy 501 TATTGAAGATCAGAGAGAGATGAGTGTCTATCTACTGCAAGCCGACCAACAGAA 560
Db 2370 TATCCGAGGTGAGGAGAGGAGTGGAGGCTCTACACCTGCCAGGCTGCAATGTCT 2429
Qy 561 GGCCTCTGTGGAAGTTTCAGCATACCTCACTGTCAAGAACTCCGACCAAGTCTAATCT 620
Db 2430 TGGCTGTGCAAGCGGAGACGCTCTTCATATAGAGGTGCCAGGAAAGACCAACTT 2489
Qy 621 GGAGCTGATCACTTAACATGCACTGTGTGGCTGGAGTCTCTTCTGGCTCTTATTAAC 680
Db 2490 GGAAGTCACTTCTGCTGCGGACTGAGTGGCATGTTCTTCTGGCTCTTCTTCTTGT 2549
Qy 681 CCTCCTTATCCGAAATGAAAGTCTC---TCTCTGAAATAGAGTCACTTACCTATC 737
Db 2550 CATTCGTACGACCCGTTAAGCGGCAATGAGGGGAACTGAGGACAGGCTACTTGTG 2609
Qy 738 AATATATGAGCAGGATGAGTTCCTTTGATGAGCAGTGTGAGCGGCTCCCTTATGA 797
Db 2610 TATGTGATGATCCAGATGAATTTGCCCTTTGGATGAGCGCTGTGAGCGCTTGCCTTATCA 2669
Qy 798 TGCACAGAGTGGAGTTCGCCGGGAGAGACTTAACTGGGCAATCACTTTGGAAGAG 857
Db 2670 TGCAGCAAGTGGGAATTCGCCGGGAGGAGGCTGAACTAGGAAACCTCTTGGCGGCG 2729
Qy 858 GGCCTTTGGAAAGTGGTTCAAGCATCAGCATTTGGCATTTGCAATTAAGAAATCACTACGTGCGG 917
Db 2730 TGGCTTGGCCCAAGTATGAGCAGAGCTTTTGGAAATGACAGACGCGCTTGTGA 2789
Qy 918 GACTGTGCTGTGAAATGCTGAAAGAGGGGGCCAGCGCCAGGAGTACAAAGCTGTAT 977
Db 2790 AACAGTAGCGCTCAAGATGTTCAAGAGAGGAGCAACACACAGCGGAGCATCGAGCCCTCAT 2849
Qy 978 GACTGAGCTAATAATCTTGACCCCATTTGCCACCATCTGACGCTGGTTACCTGTGG 1037


```

DB   3687  CC1CCTGCAAGCAAAATCGGCACGAGTATGGCAAGACTATAT 3728

RESULI 15
PCI-US92-05401-5
: Sequence 5, Application PC/TUS9205401
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: 10TPOIENT HEMATOPOIETIC STEM CELL
: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
: STREET: 180 VARICK STREET
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: U.S.A.
: ZIP: 10014
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCI/TUS92/05401
: FILING DATE: 19920626
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Feit, Irving N.
: REGISTRATION NUMBER: 28,601
: REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5405 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 208..4311
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 208..4308
PCI-US92-05401-5

Query Match 33.78; Score 708.4; DB 5; Length 5406;
Best Local Similarity 65.08; Pred. No. 4.9e-210;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps

QY 201 CACTTCATTAATCTACCATCATGAATGTTTCCTCGAAGATTCCAGGACCTATCGCTGCAG 260
DB 2070 CATCTTCATTTGGCAATTCAGAAATGCTCTCTCGAGGACCGGACGACATGTTGTC 2129
QY 261 AGCCAGGAATGATATCACAGGGGAGAAATCCTCCAGAGAAGAANAATTACAAATCAGAGA 320
DB 2130 TGCTCAAGATTAAGAAGACCAAGAAAGACATTTGGCTGGTCAACACAGCTCATCTCTAGA 2189
QY 321 TCAGSAGGACCATACCTCTCTGGCAACCTCAGTGATCACACAGATGGCCATCAGCAGCTC 380
DB 2190 GGGCATGGCACCATGATCACCGGAATCTGGAGAAATCTGGAAATTCAGACAACCAACCTG 2249
QY 381 CACCACCTTTAGACTGTATGCTAAATGGTGTCCTCCGAGCGCTCAGATCATCTGGTTTAAAAA 440
DB 2250 CATTCAGTACACTGCCAGGACTCTGGAATCCTACCCACACATTACATGTTTCAAGA 2309
QY 441 CAACCAACAATAACAAAGAGCCTGGAATTTATTTTAGGACAGGAAGCAGACGCTGTT 500
DB 2310 CAACGACACCTTGTAAGAAGATTCAAGGCAATTCAGGCAATGTTACTGAGAGATGGCAACCGAC 2369

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:50:37 ; Search time 422.586 Seconds
(without alignments)
11193.756 Million cell updates/sec

Title: x51602_COPY_1900_4000

Perfect score: 2101

Sequence: 1 TATATCAGAGTGGCCAA.....CGACAGCAGACICITGG 2101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

tal number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAI:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAI:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAI:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAI:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAI:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAI:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAI:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAI:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAI:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAI:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAI:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAI:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAI:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAI:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAI:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAI:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAI:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAI:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAI:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAI:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAI:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAI:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAI:*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAI:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2101	100.0	7680	23	AA570286
2	2097.8	99.8	4017	24	ABL91752
3	715.4	34.1	4071	20	AAV99829
4	713.8	34.0	4044	24	ABL91751
5	713.8	34.0	4071	24	ABL91754
6	713.8	34.0	4225	22	AAF83308
7	710.6	33.6	4071	13	AAQ28272
8	710.6	33.8	4236	19	AAV34763
9	708.4	33.7	5404	17	AA138735

10	708.4	33.7	5404	20	AAV77516	Murine flk-1 cDNA.
11	708.4	33.7	5406	14	AAQ53504	Murine flk-1 cDNA.
12	708.4	33.7	5406	14	AAQ35251	Human flk-1 coding
13	708.4	33.7	5406	14	AAQ40916	Murine flk-1 cDNA.
14	708.4	33.7	5406	16	AAQ81014	Flk1 receptor prot
15	708.4	33.7	5406	16	AAQ79070	Mouse flk-1 cDNA.
16	708.4	33.7	5406	18	AAV72119	Murine flk-1 recep
17	708.4	33.7	5470	15	AAQ64049	Sequence of murine
18	706.8	33.6	5406	13	AAQ29957	flk-1 cDNA sequence
19	690	32.8	2352	15	AAQ74275	truncated FLT (sVE
20	690	32.8	2352	19	AAV09333	Soluble truncated
21	690	32.8	2352	20	AAQ44435	Human soluble vasc
22	624	29.7	2523	18	AAV01457	Human VEGF recepto
23	624	29.7	2523	18	AAV52101	cDNA encoding amin
24	609.6	29.0	3009	22	AAV91131	Flt-1extrafastm/cy
25	541.6	25.8	3897	24	ABU91753	Human polynucleoti
26	541.6	25.8	4111	21	AAQ52405	Human Flt4/VEGFR-3
27	541.6	25.8	4111	22	AAQ68952	Human Flt4/VEGFR-3
28	541.6	25.8	4195	17	AAV12068	FLT4 receptor tyro
29	541.6	25.8	4195	21	AAV37815	Human Flt4 recepto
30	541.6	25.8	4195	21	AAV52333	Human tyrosine kin
31	541.6	25.8	4425	16	AAV03090	Protein tyrosine-k
32	541.6	25.8	4795	21	AAV37816	Human Flt4 recepto
33	541.6	25.8	4795	21	AAV52334	Human tyrosine kin
34	541.6	25.8	9108	16	AAV03104	Plasmid pSK5 tk1-1
35	540	25.7	4450	21	AAQ52010	Nucleotide sequenc
36	321.2	15.3	2313	19	AAV19347	Human soluble VEGF
37	321.2	15.3	2313	19	AAV09330	Soluble VEGF recep
38	321.2	15.3	2651	15	AAQ74268	SVGF-Ri gene. Ho
39	321.2	15.3	2651	20	AAQ4428	Human soluble vasc
40	321.2	15.3	7096	21	AAV75629	Nucleotide sequenc
41	321.2	15.3	7096	24	AAV39241	PD10-sflt-1 vector
42	282.8	13.5	5827	14	AAQ49753	PIK gene SAL-S1.
43	282.8	13.5	5827	15	AAV03101	Protein tyrosine-k
44	267.8	12.7	3069	21	AAZ44718	Bovine c-kit br-1
45	259	12.3	5084	19	AAV20443	Human c-kit oncoge

ALIGNMENTS

RESULT 1

AA570286
ID AA570286 standard; cDNA; 7680 BP.

AC AA570286;

XX AA570286;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #6090.

XX Human; chromosome mapping; gene mapping; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang XT;

XX WPI: 2001-632362/73.

XX P-PDS; ABQ6099.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 1: SEQ ID No 6090; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS54197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 7680 BP: 2279 A: 1661 C: 1739 G: 2001 T: 0 other:

Query Match 100.0%; Score 2101; DB 23; Length 7680;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TATATACAGATGTCGCAATGGGTTTCAGTTACTTGGAAAAATGCCGAGGA 60
 DB 1900 TATATACAGATGTCGCAATGGGTTTCAGTTACTTGGAAAAATGCCGAGGA 1959
 QY 61 GAGGACCTGAACCTGCTGCACAGTTACAGTTCATACAGACGCTTACTTGGAT 120
 DB 1960 GAGGACCTGAACCTGCTGCACAGTTACAGTTCATACAGACGCTTACTTGGAT 2019
 QY 121 TTACTGGGACAGTTAATAACAGAAATGCACTACAGTATTACCAAGCAAAAATGCC 180
 DB 2020 TTACTGGGACAGTTAATAACAGAAATGCACTACAGTATTACCAAGCAAAAATGCC 2079
 QY 181 ATCATTAGGAGACCTCCATCCTTATCTTACGATCATGATGTTCCCTGCAAGAT 240
 DB 2080 ATCATTAGGAGACCTCCATCCTTATCTTACGATCATGATGTTCCCTGCAAGAT 2139
 QY 241 TCAGGACCTATGCTTCGAGAGCCAGGAATGTATACAGAGGGAAGAAATCTCCAGAG 300
 DB 2140 TCAGGACCTATGCTTCGAGAGCCAGGAATGTATACAGAGGGAAGAAATCTCCAGAG 2199
 QY 301 AAAGAAATTAATACAGATCAGAGATCAGGACCACTACCTCTGGAAACCTCAGTATCAC 360
 DB 2200 AAAGAAATTAATACAGATCAGAGATCAGGACCACTACCTCTGGAAACCTCAGTATCAC 2259
 QY 361 ACAGTGGCCATCAGCAGTTCACCACTTACAGTCTCATCTAATGCTGCTCCGAGGCT 420
 DB 2260 ACAGTGGCCATCAGCAGTTCACCACTTACAGTCTCATCTAATGCTGCTCCGAGGCT 2319
 QY 421 CAGATCAGTCTGGTTTAAACACACCAAAATACAGAGGCTTGGAAATTAATTAGGA 480
 DB 2320 CAGATCAGTCTGGTTTAAACACACCAAAATACAGAGGCTTGGAAATTAATTAGGA 2379
 QY 481 CCAGGAACGACGCTGTTTATTGAAGATCAGAGATCAAGAGTGTCTATCAC 540
 DB 2380 CCAGGAACGACGCTGTTTATTGAAGATCAGAGATCAAGAGTGTCTATCAC 2439
 QY 541 TCGAAGCCACCAACACAGAGGCTCTGTGGAAAGTTCACATACCTACTCTTCAGGA 600
 DB 2440 TCGAAGCCACCAACACAGAGGCTCTGTGGAAAGTTCACATACCTACTCTTCAGGA 2499

QY 601 ACTCGGACAAAGTCTAATCTGGAGCTGATCCTCTAATGACCTGTGTGGCTCGGACT 660
 DB 2500 ACTCGGACAAAGTCTAATCTGGAGCTGATCCTCTAATGACCTGTGTGGCTCGGACT 2559
 QY 661 CTCCTCTGSCCTCTAATTAACCCCTCTTATCCAAAATGAAAAGGCTCTCTCTGAATA 720
 DB 2560 CTCCTCTGSCCTCTAATTAACCCCTCTTATCCAAAATGAAAAGGCTCTCTCTGAATA 2619
 QY 721 AAGNCTGACTACTCTAATTAATTAATGAGCCAGATGAAGTCTCTTGGATGAGCAGTGT 780
 DB 2620 AAGNCTGACTACTCTAATTAATTAATGAGCCAGATGAAGTCTCTTGGATGAGCAGTGT 2679
 QY 781 GAGCGGCTCCCTTATGATGCGCAGCAAGTGGGAGTTTCCCGGAGAGACTTAAACTGGGC 840
 DB 2680 GAGCGGCTCCCTTATGATGCGCAGCAAGTGGGAGTTTCCCGGAGAGACTTAAACTGGGC 2739
 QY 841 AATACACTTGAAGAGGGGCTTTTGGAAAAGTGTTCAGACATCAGCATTTGGCATTAAAG 900
 DB 2740 AATACACTTGAAGAGGGGCTTTTGGAAAAGTGTTCAGACATCAGCATTTGGCATTAAAG 2799
 QY 901 AATACACTTGAAGAGGGGCTTTTGGAAAAGTGTTCAGACATCAGCATTTGGCATTAAAG 960
 DB 2800 AATACACTTGAAGAGGGGCTTTTGGAAAAGTGTTCAGACATCAGCATTTGGCATTAAAG 2859
 QY 961 GAGTACAAAGCTCTGATGACTGAGCTTAAATAATCTTGACCCACATTTGCCACCATCTGAAC 1020
 DB 2860 GAGTACAAAGCTCTGATGACTGAGCTTAAATAATCTTGACCCACATTTGCCACCATCTGAAC 2919
 QY 1021 GTGTTTAACTGCTGGAGGCTGCACCAAGCAAGGAGGCTCTGATGCTGATTGTGAA 1080
 DB 2920 GTGTTTAACTGCTGGAGGCTGCACCAAGCAAGGAGGCTCTGATGCTGATTGTGAA 2979
 QY 1081 TACTGCAATATGAAAACTCTCCAACTACTCTCAAGAGCAAAAGTGACATTTATTTTCTC 1140
 DB 2980 TACTGCAATATGAAAACTCTCCAACTACTCTCAAGAGCAAAAGTGACATTTATTTTCTC 3039
 QY 1141 AACAGGATGACGACCTACACATGGAGCTTAAGAAAGAAATGGAGCCAGGCTGGA 1200
 DB 3040 AACAGGATGACGACCTACACATGGAGCTTAAGAAAGAAATGGAGCCAGGCTGGA 3099
 QY 1201 CAAGCAAGAAACCAAGACTAGATAGCTCACCAGCAGCAAGGCTTTTCCGAGCTCCGC 1260
 DB 3100 CAAGCAAGAAACCAAGACTAGATAGCTCACCAGCAGCAAGGCTTTTCCGAGCTCCGC 3159
 QY 1261 TTTCAGGAAGATAAAGTCTGAGTGTGATGTTGAGGAAGAGGAGGATCTGAGGGTTCTAC 1320
 DB 3160 TTTCAGGAAGATAAAGTCTGAGTGTGATGTTGAGGAAGAGGAGGATCTGAGGGTTCTAC 3219
 QY 1321 AAGGAGCCCATCAGTATGAGGATCTGATTTCTTACAGTTTTCAAGTGCCAGAGGATG 1380
 DB 3220 AAGGAGCCCATCAGTATGAGGATCTGATTTCTTACAGTTTTCAAGTGCCAGAGGATG 3279
 QY 1381 GAGTTCCTGCTCTCCAGAAATGCAATTCATCGGACCTGCGAGGAGAAATCTTTTA 1440
 DB 3280 GAGTTCCTGCTCTCCAGAAATGCAATTCATCGGACCTGCGAGGAGAAATCTTTTA 3339
 QY 1441 TCTGAGAACAAAGTGTGAGATTTGTGATTTGGCTTCCCGGGAATTTTAAAGAAC 1500
 DB 3340 TCTGAGAACAAAGTGTGAGATTTGTGATTTGGCTTCCCGGGAATTTTAAAGAAC 3399
 QY 1501 CCCGATTAATGAGAAAGAGATCTCGACTCTCTGAAATGGATGGCTCCGCAATCT 1560
 DB 3400 CCCGATTAATGAGAAAGAGATCTCGACTCTCTGAAATGGATGGCTCCGCAATCT 3459
 QY 1561 ATCTTTGAAAAATCTACAGCACCAGAGGAGCTGTGGTCTTTACCGAGTATTTGCTGG 1620
 DB 3460 ATCTTTGAAAAATCTACAGCACCAGAGGAGCTGTGGTCTTTACCGAGTATTTGCTGG 3519
 QY 1621 GAAATCTCTCTTAGTGGTCTCCATACCCAGAGTCAAAATGATGAGACTTTTTC 1680
 DB 3520 GAAATCTCTCTTAGTGGTCTCCATACCCAGAGTCAAAATGATGAGACTTTTTC 3579

Antagonists of KDR useful for treating diseases involving neovascularization, e.g. diabetic retinal vascularization, cancers (e.g. brain, breast, etc.) and forms of inflammation e.g. rheumatoid arthritis, psoriasis, contact dermatitis and hypersensitivity reactions. The polynucleotides are useful to screen for KDR antagonists/agonists and for gene therapy (e.g. by introducing a gene portion encoding a KDR protein containing functional ligand binding and membrane anchoring moieties but not tyrosine kinase activity). They are also useful to measure levels of human KDR.

11-APR-2002.
 09-JAN-2001: 2001DE-1000586.
 09-JAN-2001: 2001DE-1000586.
 (RIBO-) RIBOPHARMA AG.
 Kreutzer R, Limmer S, Rost S, Hadwiger P;
 WPI; 2002-270454/32.
 Inhibiting gene expression in cells, useful for e.g. treating tumors.
 by introducing double-stranded complementary oligoRNA having unpaired
 terminal bases
 ClaJm 13; Page 67-68; 104pp; German.
 The invention relates to a method for inhibiting expression of a target
 gene (ABL91638-ABL91797) in a cell by introducing at least one
 oligoribonucleotide that has a double-stranded structure consisting of at
 most 49 sequential nucleotide pairs, with at least part of one strand
 complementary with the target gene and has at least one end a
 single-stranded segment of 1-4 nt. The method provides
 oligoribonucleotides for antisense inhibition of gene expression useful
 e.g. for treating tumors but the oligoribonucleotides may also be
 directed against genes present in pathogens (e.g. Plasmodium or
 viruses/viroids, pathogenic on humans, animals or plants) or against
 cytokine, Id, developmental or prion genes. The method provides more
 effective inhibition of gene expression than use of known
 oligonucleotides, probably because the unpaired overhang increases
 stability and thus intracellular concentration.
 Sequence 4044 BP; 1163 A; 885 C; 1021 G; 975 T; 0 other;
 Query Match 34.0%; Score 713.8; DB 24; Length 4044;
 Best Local Similarity 65.5%; Pred. No. 2.4e-206;
 Matches 1077; Conservative 0; Mismatches 562; Indels 6; Gaps 2;
 QY 218 TCAAGATGTTCCCTGCAGATTCAGGACCTATGCTGCAGAGCCAGGAATGATACA 277
 DB 1886 TTAAGAAATGCACTCTTCAGGAGCAAGAGACATATGCTGCTTCAGAGCAAGA 1945
 QY 278 CAGGGAGAAATCCTCCAGANAGAAATTAACAATCAGATCAGAGCAGGACCAATACC 337
 DB 1946 CAGAGAAAGACATGCTGCTCAGCAGCTCAGAGCTCAGAGCTGAGGACCCAGCA 2005
 338 TCTCGAAACCTCAGTATCAGACAGTGGCCATCAGCAGTTCACCACTTTAGACTGTC 397
 2006 TCACAGGAACCTGGAAATCAGACAGCAAGATGTTGGGAAAGCATCGAATCTCATGCA 2065
 QY 398 ATGCTAATGGTGTCCCGAGCTCAGATCAGTCTGTTTAAACACACCAAAATACAAAC 457
 DB 2066 CGGCATCTGGGAATCCCGCTCCACATCATGTTGGTTTAAAGATAATGAGACCTTTAG 2125
 QY 458 AAGAGCTGGAAATATTATAGGACCAAGAGCAGCAGCTTTTATGAAGAGTCAGAC 517
 DB 2126 AAGACTCAGGCAATGATAGAGGTTGGAAACCGGAACTCCTATCCGAGAGTGA 2185
 QY 518 AAGAGATGAGGTGCTATCAGTCAAGAGCCACCAAGGAGGCTGTGGAAGTT 577
 DB 2186 AAGAGACAGAGGCTCTACACCTCGAGGCAATGAGTCTTGGCTGTGCAAAAGTGG 2245
 QY 578 CAGCATACCTCAGTCTCAGAGACCTCGGACAGTCTAATCTGGAGCTGATCACTTAA 637
 DB 2246 AGGCATTTTCAATAGAGAGTGGCCAGGAAAGACCACTTGGAAATCATTTATCTAG 2305
 QY 638 CATGCACCTGTGTGCTGCGACTCTTCTGCTCTTATTAACCTCTTATCCGAAAAA 697
 DB 2306 TAGGACGCGGTGATGCCAATGTTCTTCUGGCTACTTCTGCTATCACTACGACCG 2365
 QY 698 TGAAGAGTCTCTTCTTCTGAAATAAGAGACTGACTACTAATTAATATGAGCCAG 754

DB 2366 TTAAGCGGCAATGGAGGGAATCAAGACAGGCTACTTGTCCATGTCATGATCAG 2425
 QY 755 ATGAAGTTCCTTTGGATGAGCAGTGTAGCGGCTCCCTTATGATGCCAGAGTGGAGT 814
 DB 2426 ATGAATCCCAATGGATGAACATTTGTAACGACTGCTTATGATGCCAGCAATGGGAAT 2485
 QY 815 TTSCCGGGAGAGACTTAACTGGGCAATCAGCTTGAAGAGGGGCTTTTGAAGTGG 874
 DB 2486 TCCCGAGAGACCGGCTGAGCTAGGTAAAGCTTTGGCGGTGGTGGCTTTGGCCAAGTGA 2545
 QY 875 TTCAAGCAITACAGATTTGGCATTAAGAATCAGCTGCGGAGCTGCGGAGTGTGCTGTA 934
 DB 2546 TTGAGCAGATGCTTTTGAATTTGAACAGACAGCAACTTTCAGGACAGTAGCAGTCAAAA 2605
 QY 935 TGTCAAGAGAGGGGCGGCGGCGAGCTCAAAAGCTCTGATGACTGAGCTTAAATATCT 994
 DB 2606 TGTCAAGAGAGGAGCAACACAGCTGAGCATCGAGCTCTCATGCTGAGCTCAAGATCC 2665
 QY 995 TGACCCCAATTTGGCCACCATCTGAGCTGGTTAAGCTTGTGGAGGCTGCGACCAAGCAAG 1054
 DB 2666 ICATTCATATTTGGTCAACCATCTCAATGTTGGTCAACCTTCTAGGCTGCTGACCAAGCAG 2725
 QY 1055 GAGGCGCTCTGATGCTGATTTGTAATCTGAAATATGAAATCTCTCCAACTACCTCA 1114
 DB 2726 GAGGCGCACTCATGCTGATTTGTAATCTGAAATTTGAAACCTTCTCCACTTACCTGA 2785
 QY 1115 AGAGCAACGCTGACTTTATTTTCTCAAGAGGATGAGCAGCTACACATGAGGAGCTAAGA 1174
 DB 2786 GGAGCAAGAGAAATGAATTTGCTCCCTACAGACCAAGGGCGACGATTTCCGTCAGGGA 2845
 QY 1175 AAGAAATATGGAGCGGCTGACAGGCTGACAGAGCAAGCAACAGCACTAGATAGCTACCA 1234
 DB 2846 AGACTAGCTTGGAGCAATCCCTG---TGGATCTGAACGGGCTTGGACAGCATCAGCA 2902
 QY 1235 GCAGCAAAAGCTTTCCGAGCTCCGCTTTTACAGAGATATAAAGTCTGAGTATGTTGAGG 1294
 DB 2903 GTAGCCAGAGCTCAGCCAGCTCTGGATTTGTGGAGAGAGTCCCTCAGTGTAGTAGAG 2962
 QY 1295 AGAGGAGATTTGAGCGTTTCTACAGAGGCGCATCAGTATGGAAGATCTGATTTCTT 1354
 DB 2963 AAGAGAAAGCTCTCTGAAGATCTGTATAAGGACTTCTTGAGCTTGGAGCATCTCATCTGT 3022
 QY 1355 ACAGTTTCAAGTGCAGAGGAGTGGAGTCTCTTCTCCAGAAAGTCAATCATGCGG 1414
 DB 3023 ACAGTTTCAAGTGCCTAGGAGTGGAGTCTTGGCAATCGGAAAGTATATCCAGAGG 3082
 QY 1415 ACCTGCGAGAGAGAAATCTTTTATCTGAGAAACAGCTGGTGAAGATTTGATTTTG 1474
 DB 3083 ACCTGCGGCGAGAAATATCTCTTATCGGAGAGAGCTGGTTTAAATCTGTGACTTTG 3142
 QY 1475 GCCTTCCCGGGATTTTATAGAACCCGATTTATGTAAGAGGAGTATCTCGACTTC 1534
 DB 3143 GCTTGGCGGGATTTTATAGATCCAGATTTATGTAAGAGAGAGATGCTGCCCTCC 3202
 QY 1535 CTCTCAATGGATGGCTCCGAAATCTATCTTTGACAAATCTACACCAACCAAGAGCGAG 1594
 DB 3203 CTTTGAAATGGATGGCCCGGAGAAACAAATTTTGTACAGAGTGTACACAAATCCAGAGTGA 3262
 QY 1595 TGTGCTTTAGGAGTATGCTGTGGAAATCTCTCTCTAGTGGTCTCTCATACCCAG 1654
 DB 3263 TCTGCTTTTGGTGTCTGCTGTGGAAATTTTCTCTAGTGGTCTCTCATATCTG 3322
 QY 1655 GAGTCAATATGAGGAGCTTTTGTGAGTCCGCTGAGGAGGAGTGAAGAGTGAAGCTC 1714
 DB 3323 GGTAAAGATGATGAGATTTTGTAGCGGATTAAGAGAGGAGTGAAGTGAAGGCGCC 3382
 QY 1715 CTGAGTACTTACTCTTGAATCTATCAGATCATCTGAGCTGCTGCGACAGAGACCA 1774
 DB 3383 CTGATTTATACACCAAGAAATGTACAGACCATCTGCTGAGTGTGCGCAGGAGCCCA 3442
 QY 1775 AAGAAAGGCGAGATTTGCGAGACTTGTGGAATAACTAGTGTATTTGCTTCAAGCAAT 1834

Db 3443 GTGAGAGACCCAGCTTTTACAGAGTGTGGAGCAATTCGGAAATCTCTGCAAGCTAATG 3502
QY 1835 TACAACAGAGTGTAAAGACTACAT 1859
Db 3503 CTCAGCAGGATGCAAGACTACAT 3527
RESULT 5
ABL91754
ID ABL91754 standard: DNA: 4071 BP.
XX
AC ABL91754;
XX
DI 28-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 97.
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW plasmidium; virus; viroid; cytokine; prion; antisense oligonucleotide;
CYTOSTATIC; virucide; protozoacide; antibacterial; ds.
XX Homo sapiens.
XX
PN DE10100586-C1.
XX
PD 11-APR-2002.
XX
PF 09-JAN-2001; 2001DE-1000586.
XX
PR 09-JAN-2001; 2001DE-1000586.
XX
PA (RIBO-) RIBOPHARVA AG.
XX
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX WPI; 2002-270454/32.
XX
PT Inhibiting gene expression in cells, useful for e.g. treating tumors,
PT by introducing double-stranded complementary oligoRNA having unpaired
PI terminal bases
XX
PS Claim 13; Page 71-72; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-ABL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a
CC single-stranded segment of 1-4 nt. The method provides
CC oligoribonucleotides for antisense inhibition of gene expression useful
CC e.g. for treating tumours but the oligoribonucleotides may also be
CC directed against genes present in pathogens (e.g. plasmidium or
CC viruses/viroids, pathogenic on humans, animals or plants) or against
CC cytokine, id, developmental or prion genes. The method provides more
CC effective inhibition of gene expression than use of known
CC oligonucleotides, probably because the unpaired overhang increases
CC stability and thus intracellular concentration.
XX
SQ Sequence 4071 BP; 1169 A; 894 C; 1025 G; 993 T; 0 other;
Query Match 34.0%; Score 713.8; DB 24; Length 4071;
Best Local Similarity 55.5%; Pred. No. 2.4e-206;
Matches 1077; Conservative 0; Mismatches 562; Indels 5; Gaps 2;
QY 218 TCATGATGTTTCCCTGCAAGATTCAGGACCTATGCTGCGAGAGCCGGAATGTATACA 277
Db 1886 TTAAGATGATCTCTTGCAGGACCAAGAGACTATGCTGCTTCTCTCAAGACAGGAAGA 1945
QY 278 CAGGGGAAGAAATCTCTCCAGGAAGAAGAAATACATACAGATCAGAGACAGACCAATCC 337
Db 1946 CCAGAAAGACATTCCTGGTGTAGGACGCTACAGTCTCTGAGCGGTGTGGACCCACGA 2005
QY 338 TCCTGGAAACCTCAGTGTATCACACAGTGGCCCAICAGCAGTTCACCACCTTATAGACIGTC 397

Db 2006 ICACAGAAACCTTGAGATCAGAGCAAGTATTCGGGAAGCATCGAGTCTCATGA 2065
QY 398 ATGCTAATGTTGCTCCGAGGCTCAGATCACTTGGTTTAAAAACACCAAAATACAC 457
Db 2066 CGCATCTGGGAATCCCTCTCCAGATCATGTGGTTTAAAGATAANTGAGACCTTGTAG 2125
QY 458 AAGAGCTGSAATTAATTTAGGACCAAGCAGCAGCTGTTTATTCAAGAGATCACAG 517
Db 2126 AAGACTCAGGCAATGATGAGAGATGGAAACCGGAACCTACTATCCGAGAGTGA 2185
QY 518 AAGAGGATGAGTGTCTATCACTGCAAGCCACCAAGGGCTCTCTGGAAAGTT 577
Db 2186 AAGAGGACGAAGGCTCTACACCTGCCAGCATGCAGTGTCTTGGCTGTGCAAGTGS 2245
QY 578 CAGCATACCTCACTGTTCAGGAACCTCGCAGCAAGTCTTAATCTGAGCTGATCACTCTAA 637
Db 2246 AGCATTTTTCATAATGAAGTGCAGGAAAGAGCAAGCTTGGAAATCATTTATCTAG 2305
QY 638 CATGCCACTGTGTGGCTCGGACTCTCTTCTGGCTCTTAATTAACCTCTCTTATCCGAAAA 697
Db 2306 TAGCAGCGCGGTATGCCATGTCTTCTGGCTACTTCTGTCAATCATCTACGAGCG 2365
QY 698 TGAAGAGTCT---TCTTCTGAAATAAGACTGACTTACATTAATTAATGACCCAG 754
Db 2366 TTAAGCGGCAATGAGAGGGAAGTGAAGAGAGGCTACTTGTCCATCTCATGATCCAG 2425
QY 755 ATGAAGTCTCTTGGATGAGGAGTGTGAGCGCTCCCTTATGATGCCAGCAAGTGGAGT 814
Db 2426 ATGAATCTCCATGATGAACATTTGTAAGCAGCTGCCCTTATGATGCCAGCAATGGGAAT 2485
QY 815 TTGCCGGGAGAGACTTAACTGGGCAATCACTTGGAAAGAGGGGCTTTTGGAAAAAGTGG 874
Db 2486 TCCCAAGAGACCGGCTGAAGCTAGGTAAAGCTCTTGGCCGTGGTGGCTTTGGCCAGTGA 2545
QY 875 TTCAAGCATCAGCAATTTGSCATTAGAAATCACTTACCTAGCTGCGGAGCTGTGGTGTGAAA 934
Db 2546 TTGAAGCAATGCTTTTGAATTTGCAAGACAGCAACTTTCAGGACAGATGACAGTCAAAA 2605
QY 935 TGTCTAAAGAGGGGCGCCAGCGAGGTACAAAGCTCTGATGACTGAGCTAAATCT 994
Db 2506 TGTGAAGAGAGGACACACAGTACAGTCTGAGCTCTCATGTCTGAATCAAGATCC 2665
QY 995 TGACCCAAATGGCCACCATCTGAACCTGTTTAACTGCTGGAGCTTGCACCAAGCAAG 1054
Db 2566 TCAATTCATATGCTCACCCTCAATGCTGCTAGCTTCTAGGTGCTGTACCAAGCCAG 2725
QY 1055 GAGGGCTCTGATGGTGTGATGCTGAATCTGCAATATGGAATCTCTCCAACTACTCA 1114
Db 2726 GAGGGCCACTCAATGGTGTGATTTGGAATCTTGGCAATTTGGAAACCTGTCCACTTACCTGA 2785
QY 1115 AGAGCAAGCTGACTTATTTTTCTCAAGAGTGCAGCACTACACATGGAGCTTAAGA 1174
Db 2786 GGAGCAAGAGAAATGAATTTGTCCTCTACAAGACCAAGGGGCGACGATTCGTCAGGGA 2845
QY 1175 AAGAAAAATGAGCGAGCGCTGCAAGAGGCAAGAAACCAAGACTAGTAGGCTCACCA 1234
Db 2846 AAGACTACCTTGGAGCAATCCCTG---TGGATCTGAAGCGCGCTTGGACAGCATCACCA 2902
QY 1235 GCAGCAAGAGCTTTTCCGAGCTCCGGCTTTTCCAGCAAGATAAAAGTCTGAGTGTGTTGAG 1294
Db 2903 GTAGCCAGAGCTCAGCCAGCTCTGGATTTGTGGAGGAGAGTCCCTCAGTGTATGAGAG 2962
QY 1295 AAGAGGAGATTTCTGACGGTTTCTACAAGAGGCCCATCACTATGGAAGATCTGATTTCTT 1354
Db 2963 AAGAGGAAGCTCTCAAGATCTGTATAAGGACTTCTGACCTTGGAGCATCTCATCTGTT 3022
QY 1355 ACAGTTTCAAGTGGCCAGAGGATGGAGTTCTGTCTTCCAGAAAGTGCATTCATCTCGG 1414
Db 3023 ACAGTTTCAAGTGGCCAGAGGATGGAGTTCTGTGCAATTCGCAATTCGAGAGTGTATCCACAGG 3082
QY 1415 ACCTGGCAGGAGAAACATCTTTTATCTGAGAAACAGTGTGAGATTTGTGATTTTG 1474

Db 3363 CTGATTAATACACACAGAAATGTACCAACAGCCATGCTGGAGCTGCTGGACAGGGGAGCCCA 3444

QY 1775 ANGAAGGGCCAGAAATTTGCAAGAACTTGTGGAAAACCTAGTGAATTCCTCAAGCAAAATG 1834

Db 3443 CTCAGAGACCCACGTTTTCAGAGTTGGTGACAAATTGGGAAAATCTCTTCAAGACTAATG 3502

QY 1835 TACACACAGATGGTAAGACTACAT 1859

Db 3503 CTCAGCAGGATGGCAAGACTACAT 3527

RESULT 8

AAV34763

ID: AAV34763 standard; DNA: 4236 BP.

XX AC AAV34763:

XX

XX 27-AUG-1998 (first entry)

XX Human KDR genomic DNA.

XX Kinase insert domain containing receptor; KDR; screening; inhibitor;

XX vascular endothelial growth factor; VEGF; angiogenesis; treatment;

KW cancer; ss.

XX

XX Homo sapiens.

OS

XX

XX Key Location/Qualifiers

FI 1..4071

FI CDS

FI /tag= a

FI /product= KDR

FI /note= "kinase insert domain containing receptor"

XX

XX US5766860-A.

XX

XX 16-JUN-1998.

XX

XX 25-FEB-1997; 97US-0810116.

XX

XX 23-NOV-1992; 92US-0930548.

XX 25-FEB-1997; 97US-0810116.

XX (AMCY) AMERICAN CYANAMID CO.

XX

XX Carrión ME, Terman BI.

XX

XX WPI: 1998-361682/31.

XX P-PSDB: AAV59275.

XX

XX Screening assay for vascular endothelial cell growth factor

PI antagonists - using recombinant cells expressing receptor protein

XX

XX Claim 1; Fig 7A-M; 51pp; English.

XX

XX This sequence encodes a novel human growth factor receptor, kinase

XX insert domain containing receptor or KDR. This receptor is capable of

XX binding to the vascular endothelial cell growth factor, VEGF and is

XX used in a screening assay which identifies compounds that inhibit VEGF

XX action on KDR. Such compounds which inhibit binding of VEGF to the KDR

XX may inhibit angiogenesis and thus be useful for treating cancer.

XX

XX Sequence 4236 BP; 1215 A; 938 C; 1062 G; 1020 T; 0 other;

QY

Query Match 33.8%; Score 710.6; DB 19; Length 4236;

Best Local Similarity 65.3%; Pred. No. 2.4e-205;

Matches 1075; Conservative 0; Mismatches 564; Indels 6; Gaps

QY 218 TCATGAATGTTCCCGCAGATTCAGGCACCTATGCTTCGACAGCCAGGAATGTATACA 277

Db 1895 ITAAGAATGCACTTCGAGGACCAAGGAGACTATGTCGTGCTTCTCTCAAGACAGGAAGA 1945

QY 278 CAGCGGAGAAATCTCCAGAGAAAGAAATTAACAATACAGATCAGGACGACCAATCC 337

Db 1945 CCAGAAAGACATTCGGTGGTCAGGCAGCTACAGTCTTACAGGCTGTGCGCCACGCA 2005
 QY 338 TCTTCCGAAACCTCAGTGATACACAGTGGCCATCAGCAGTTCACACCTTACAGCTC 397
 Db 2005 TCACAGGAACCTCGGAAATCAGACACAGTATGGGGAAGCATCGAGTCTCATGCA 2065
 QY 398 ATGCTAATGTTGTCGCGAGCTCAGATCAGTCTGGTTTAAACACACCAAAATACAC 457
 Db 2065 CGCATGTCGGAAATCCCTCCACAGATCATGTGTTTAAAGATATGAGACCTTGTAG 2125
 QY 458 AAGACCTTGGAAATTTATAGCAGCAGCAAGCAGCAGCTGTTTATGAAAGATCAG 517
 Db 2125 AAGACTCAGCAATGATGAGGATGGACCGGACCTCCTATCCGAGATGAGGA 2185
 QY 518 AAGAGATGAGGTGCTTACTTCCAAAGCCCAACACAGAGGCTCTGTGGAAGAT 577
 Db 2195 AGCAGGAGGAGGCTCTACCTCCAGGATGCTGTTTAAAGATATGAGACCTTGTAG 2245
 QY 578 CAGTACTCTCAGTGTTCAGCAAGCTCGGACAGCTTAACTGTGAGGTGATCCTNA 637
 Db 2245 AGGCAATTTTCAATATAGAGTGGCCAGGAAAGAGCACTTGGAAATCATATTTAG 2305
 QY 638 CATGACCTGTGGTGGGACTCTCTTCTGCTCTTATTAACCTCTTATCCGAAAAA 697
 Db 2305 TAGGCAGCAGGTGATGCTGCTCTTCTGCTTACTTCTGCTATCTCTAGGACCG 2365
 QY 698 TGAAGATC---TCTCTGAAATTAAGACTGACTACCTATCAATATATGAGCCAG 754
 Db 2365 TTAAGCGGCAATGAGGAGGCACTGAAGACAGGCTACTTGTCCATCGTATGATCAG 2425
 QY 755 ATGAAGTCTCTTGGATGAGGAGTGTGAGGCTCCCTTAATGATGCGCAGGATGGAGT 814
 Db 2425 ATGAATCTCCATGGATGAACTGTGAGCAGCTGCCCTTAATGATGCGCAGGAAAT 2485
 QY 815 TTGCGCGGAGAGACTTAACTGGGCAATCATTGGAAGAGGGCTTTTGGAAAAATG 874
 Db 2486 TCCCGAGAGCGGCTGAACCTAGGTAGGCTCTTGGCGGTGGCTTGGCCAGAGA 2545
 QY 875 TTCAAGATCAGCAATTTGGCAATAGAAATCAGCTAGTCCGAGCTGTGGCTGTGAAA 934
 Db 2546 TTGAAGCAGATGCTTTGGAATTTGCAAGACAGCACTTGCAGGACAGTACAGCTCAA 2605
 QY 935 TGCTGAAGAGGGGCGCGGCGGAGTACAAAGCTCTGATGACTAGCTTAAATATCT 994
 Db 2605 TGTTGAAGAGGAGCAACACAGAGTACGAGCTCTGATGCTGAGCTCAAGATCC 2665
 QY 995 TGACCCATATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1054
 Db 2665 TCAATCATATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2725
 QY 1055 GAGGCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1114
 Db 2726 GAGGCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2785
 QY 1115 AGAGCAACGTGACTTATTTTCTCAAGAGTACAGCTACAGCTACAGCTACAGCTACAG 1174
 Db 2786 GGAGCAACAGAAATGAAATTTGCTCAAGAGTACAGCTACAGCTACAGCTACAGCTAC 2845
 QY 1175 AGAGCAACGTGACTTATTTTCTCAAGAGTACAGCTACAGCTACAGCTACAGCTACAG 1234
 Db 2846 AAGACTGCTTGGCAATCCCTG---TGAATCTGAAGAGGCTTGGCAATCCCTG 2902
 QY 1235 CGAGCAAGCTTGGCAATCCCTG---TGAATCTGAAGAGGCTTGGCAATCCCTG 1294
 Db 2903 GTAGCCAGCTCAGCCAGCTTGGCAATCCCTG---TGAATCTGAAGAGGCTTGGCAAT 2962
 QY 1295 AAGAGGAGTCTTGGCAATCCCTG---TGAATCTGAAGAGGCTTGGCAATCCCTG 1354
 Db 2963 AAGAGGAGTCTTGGCAATCCCTG---TGAATCTGAAGAGGCTTGGCAATCCCTG 3022
 QY 1355 ACAGTCTTGGCAATCCCTG---TGAATCTGAAGAGGCTTGGCAATCCCTG 1414
 Db 3023 ACAGTCTTGGCAATCCCTG---TGAATCTGAAGAGGCTTGGCAATCCCTG 3082

QY 1415 ACCTGGAGCGAGAAACATTTCTTATCTGAGAAACACAGTGTGAGATTTGTGATTTG 1474
 Db 3083 ACCTGGAGCGAGAAACATTTCTTATCTGAGAAACACAGTGTGAGATTTGTGATTTG 3142
 QY 1475 GCCTGGCGGAGATTTATTAAGAACCCCGATTTATGTGAGAAAGAGATATCTGATTT 1534
 Db 3143 GCCTGGCGGAGATTTATTAAGAACCCCGATTTATGTGAGAAAGAGATATCTGATTT 3202
 QY 1535 CTCTGAATGAGTGTGCTGCGGATCTTCTTGTGAGAAATCTACAGAACTACAGCAGGACG 1594
 Db 3203 CTCTGAATGAGTGTGCTGCGGATCTTCTTGTGAGAAATCTACAGAACTACAGCAGG 3262
 QY 1595 TGTGCTTACGAGATTTCTGCTGAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAAT 1654
 Db 3263 TGTGCTTACGAGATTTCTGCTGAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAAT 3322
 QY 1655 GAGTACAAATGAGTGTGCTGCGGATCTTCTTGTGAGAAATCTTCTTGTGAGAAAT 1714
 Db 3323 GAGTACAAATGAGTGTGCTGCGGATCTTCTTGTGAGAAATCTTCTTGTGAGAAAT 3382
 QY 1715 CTGAGTACTTCTGCTGAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAATCTTCT 1774
 Db 3383 CTGATTAATCTACAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAATCTTCT 3442
 QY 1775 AAGAAAGCGGAGATTTGAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAAT 1834
 Db 3443 GTGAGAAAGCGGAGATTTGAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAAT 3502
 QY 1835 TACAACAGGATGTGAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAATCTTCT 1899
 Db 3503 CTGAGAGATGTGAGAAATCTTCTTGTGAGAAATCTTCTTGTGAGAAATCTTCT 3574

RESULT 9

AAT38735
 ID AAT38735 standard; cDNA; 5404 BP.

XX AC AAT38735;

XX DT 11-DEC-1996 (first entry)

XX DE Murine foetal liver kinase 1 cDNA.

KW Marine: foetal liver kinase 1; fik-1; protein tyrosine kinase;
 KW monoclonal; antibody; extracellular domain; receptor assay;
 KW haematopoietic stem cell; ligand; stimulation; proliferation;
 KW differentiation; treatment; anaemia; bone marrow damage;
 KW cancer chemotherapy; radiation; ds.

XX CS Mus musculus.

XX FH Key Location/Qualifiers

FI CDS 208..4311

FI sig_peptide 208..264

FI tag- b 265..4308

FI tag- c 265..4308

XX FN CS5548065-A.

XX XX 20-AUG-1996.

XX XX 02-APR-1991; 91US-0679666.

XX XX 19-NOV-1992; 92US-0977451.

XX XX 02-APR-1991; 91US-0679666.

XX XX 28-JUN-1991; 91US-0728913.

XX XX 15-NOV-1991; 91US-0793065.

XX XX 24-DEC-1991; 91US-0813593.

XX XX 26-JUN-1992; 92US-0906397.

XX XX 12-NOV-1992; 92US-0975049.

PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252517.
 XX
 PA (UYP-1) UNIV PRINCETON.
 XX
 XX Lemischka IP;
 XX
 DR WPI; 1996-392678/39.
 DR P-PSDB; AAT38735.
 XX
 XX
 PT Anti-fetal liver kinase 2 (flk-2) antibodies - useful in assays,
 PT for isolating haematopoietic stem cells expressing receptor and for
 PT obtaining ligands
 XX
 XX
 PS Disclosure: Columns 51-62; 50pp; English.
 XX
 CC The present sequence encodes murine foetal liver kinase 1 (flk-1),
 CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
 CC raised against the extracellular portion of flk-1 can be used to
 CC assay for flk receptors on the surface of haematopoietic stem
 CC cells, and to isolate positive cells. The antibodies can also
 CC be used as, or to obtain ligands, which stimulate the proliferation
 CC and/or differentiation of stem cells. The ligands can be used, e.g.
 CC for treating anaemia, or bone marrow damage resulting from cancer
 CC chemotherapy, or radiation.
 XX
 XX Sequence 5404 BP: 1411 A; 1297 C; 1423 G; 1273 T; 0 other:
 SQ
 Query Match 33.7%; Score 708.4; DB 17; Length 5404;
 Best Local Similarity 65.0%; Pred. No. 1.3e-204;
 Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;
 QY 201 CACTTATATCTACCATCATGATGTTCTCCGTCGAAGATTCAGGCACCTATGCGTGCAG 260
 DB 2070 CATCTGATGTTGGCAITTCAGATGCTCTGCGAGGCAAGGCGACTATGTTGCTC 2129
 QY 261 AGCCAGGAATGTATACACAGGCGGAAGAAATCTCCAGAGAAAGAAATTAACAATCAGAGA 320
 DB 2130 TGCTCAAGATAGAGACCAAGAAAGACATGCTTGTCACACAGCTCATCACTCTGA 2189
 QY 321 TCAGGAGACCAATACCTCTCGGAACCTCAGTATCACACAGTGGCCATCAGAGTTC 380
 DB 2190 GCGCATGGCACCCATGATTCACCGGAATCTGGAGAAATCAGACAAACCAATGGCGAGAC 2249
 QY 381 CACCACCTTTAGACTGTCAIGCTAATGTGTCCCGAGGCTCAGATCATCTTGTTTAAAAA 440
 DB 2250 CATTGAGTGACTTGGCCAGCATCTGMAATCTACCCCAACATATACATGGTTCAGAGA 2309
 QY 441 CAACCAAAAATACAAAGAGCCTGGAATATTTTAGACCAAGAGCAGCAGCTGTT 500
 DB 2310 CAACGAGACCCCTGGTAGAAGATTCAGGCATTTGACTGAGAGATGGGAACCGGAACCTGAC 2359
 QY 501 TATTGAAGAGTACAGAGAGGATGAAGGTGCTATCACTGCAAGGCCACCAACCAAGAA 560
 DB 2370 TATCCGAGGTTGAGGAGGAGGATGAGGCGCTCTACCTGCCAGGCGCTGCAATGCT 2429
 QY 561 GGCTCTGTGGAAAGTTCAGATACCTCAGTCTGTTCAAGGACCTCGGACAAGCTAATCT 620
 DB 2430 TGGCTGTGCAAGCGGAGAGCGCTCTTCATATAGAGGTGCCAGGAAAGACCAACTT 2489
 QY 621 GGAGCTGATCACTCTACATGCACCTGTGGCTGGGACTCTCTCTGGCTCTCTATTAAC 680
 DB 2490 GGAAGTCAATATCTCTGCGCACTGCAGTGAITGCCATGTTCTCTGGCTCTCTTGT 2549
 QY 681 CCTCTTATCCGAAAATGAAGAGTC-...TCTCTGAAATGAAGACTGACTACCTATC 737
 DB 2550 CATCTCTGACGAGCCGTTTACGGGGCCAAAGAGGGGAACTGAGACAGGCTACTTGTC 2609
 QY 738 AATTATATGGACCCAGATGAAGTCTCTTTGGATGAGCAGTGTGAGCGGCTCTTATGA 797
 DB 2610 TATGTCAATGGATCCAGATGAATGCCCCITGGATGAGCGCTGTGACGCTGCTTATGA 2663
 QY 798 TGCCAGCAAGTGGGAGTTTGCCCGGAGAGACTTAACTGGGCAATCATCTGGAGAG 857

RESULT 10

DB 2570 TGCCAGCAAGTGGGAATTCGCCAGGACCGGCTGAACACTAGAAAAACCTCTTTGGCGCGG 2729
 QY 858 GCGTTTGGAAAAGTGGTTCAAGCATCAGCATTTTGGCAATTAAGAATACACCTACCTGCGCG 917
 DB 2730 TGGCTTCGGCAAGTATTGAGCGCACCGCTTTTGGAAATTCACACAGCAGCAGCTTGC 2789
 QY 918 GACTGTGGCTGTGAATAATGCTGAAAGAGGGGCCACGGCCACGGAGTAGTACAAAGCTCT 977
 DB 2790 AACAGTAGCGGTCAAGATGTTGAAGAAGGAGCAACACACAGCGAGCATCGAGCCCTCA 2849
 QY 978 GACTGAGCTTAAATCTTGACCCACATTCGCCACCATCTGACCTGCTTAACTCTCTGG 1037
 DB 2850 GTCTGAATCAAGATCTCTATCCACATCTGTCACCATCTCAATGTGGTGAACCTCTAGG 2909
 QY 1038 AGCTGCACCAAGAGAGGGCTCTGATGGTGAATGTTGAATACTGCAATATGGAAA 1097
 DB 2910 CGCTTCACCAAGCGGAGGGCTCTCATGTGATGTTGGAAATCTCGAATTTGGAAA 2969
 QY 1098 TCTCTCCACTACCTCAAGAGCAAGCTGACATTTTCTCAACAGGATGCAGACT 1157
 DB 2970 CCTATCAACTTACTTACGGGCAAGAGAAATGAATTTGTTCCCTATAAGAGCAAAAGG 3029
 QY 1158 ACATGAGGAGCTTAAGAAAAGAAATGGAGCCAGGCTCGAACAAGCAAGAACCAAG 1217
 DB 3030 ACGTTCGCCAGGCAAGGACTACGTTGG--GAGCTCTCCGTGGATCTGAAGAGC 3086
 QY 1218 ACTAGTACGCTACACAGCAGGCAAGCTTTGCGAGCTCCGGCTTTCAGGAGATAAAG 1277
 DB 3087 CTGGACAGCATCACACAGCAGGAGCTCTGCCAGCTCAGGCTTTGTTGAGGAGAAATC 3146
 QY 1278 TCTGAGTGAATGTTGAGGAGAGGAGATCTCAGGCTTTCTACAGGAGCCCATCAT 1337
 DB 3147 GCTCAGTGAATGAGGAGAGGAGCTCTGAGAACTGTACAGGAGCTTCTGACCT 3206
 QY 1338 GGAAGATCTGATTTTACAGTTTCAAGTGGCCAGGCAATGAGTTCCTGTCTTCCAG 1397
 DB 3207 GGAGCATCTCTCTTACAGCTTCCAAGTGGCTAAGGCAATGGAGTCTTGGCATCAAG 3266
 QY 1398 AAGTGCATTCATCGGACCTGCGAGCGGAGAACTCTTTATCTGAGAACAACTGGT 1457
 DB 3267 GAAGTGTATCCAGGGAGCTTGGCAGCAGCAAACTCTCTCTCGGAGAGAAATGTTGT 3326
 QY 1458 GAAGATTTGTGATTTTGGCTTGGCGGGATTTTAAAGAACCCCGATTTATGTGAGAAA 1517
 DB 3327 TAAGATCTGAGCTTCGCTTTGGCCGGGACATTTAAAGACCCGGATTTATGTGAGAAA 3386
 QY 1518 AGGAGATCTCGACTTCTCTGAAATGGATGGCTCCGAACTCTATCTTTGACAAATCTA 1577
 DB 3387 AGGAGATCGCCCACTCCCTTTGAAGTGGATGGCCCGGAAACCAITTTTTCACAGAGTA 3446
 QY 1578 CAGCAACAAGCAGCAGCTGTGCTTACGGAGTATTGTGTGGGAAATCTTCTCTTAGG 1637
 DB 3447 CACAATTCAGACGGAATGTGGTCTTTCGGTGTGTCTCTGGGAAATATTTTCTTAGG 3506
 QY 1638 TGGGTCTCCATACCAGAGGTACAAATGGATGAGGACTTTTGCAGTCCGCTGAGGAGAG 1697
 DB 3507 TGGTCTCCCATACCCTGGGGTCAAGATTTGATGAAGAAATTTGTAGGAGATTGGAAGAGG 3566
 QY 1698 CATGAGGATCAGAGCTCTGAGTACTCTACTCTGAAATCTATCAGATCATCTGAGACTG 1757
 DB 3567 AACTGAATGCGGGGCTCTCTACTACTACCTCCAGAAATGTACCAGACCAATGCTGGAGTG 3626
 QY 1758 CTGSCACAGAGACCCCAAGAAAGAGCCAGATTTGCAAGCTTGTGGAAAACCTAGGTGA 1817
 DB 3627 CTGGATGAGGACCCCAACAGAGACCCCTCGTTTTCAGAGTTGGTGGAGCATTTGGGAAA 3686
 QY 1818 TTTGCTTCAAGCAATGTACAAAGAGGATGGTAAAGACTACAT 1859
 DB 3687 CCTCTTGAAGCAATGCGCAGCAGGATGGCAAGACTATAT 3728

[illegible]

Db 2790 AACAGTAGCGGTCAAGATGTTGAAGAGAGGAGCAACACACACACGAGCAICAGGCCICAT 2849
Qy 978 GAGTGAAGCTAAATAATCTTACACCCACATTCGACACCATCTGACAGTGTGTTAACTGTGTGG 1037
Db 2850 GTCTGAATCAAGATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCA 2909
Qy 1038 AGCTGACACAG 1097
Db 2910 GCGCTGACACAAACCGGAG 2969
Qy 1038 TCTCTCACTACCTACACAG 1157
Db 2970 CCTATCAACTTCTTACGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3029
Qy 1158 ACACATGGAGCCTTAAG 1217
Db 3030 ACGCTTCGCGAG 3086
Qy 1218 ACTAGATAGCGTACACAG 1277
Db 3087 CTGAG 3146
Qy 1278 TCTGAGTATGTTAGAT 1337
Db 3147 CCTCAGTATGAT 1396
Qy 1338 GGAAGATCTGATTTTACAGTTTCAAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1397
Db 3207 GAGAGATCTGATTTTACAGTTTCAAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3266
Qy 1398 AAGATGATCTGAT 1457
Db 3267 GAGAGATCTGAT 1516
Qy 1458 GAGAGATCTGAT 1517
Db 3327 TAAGATCTGAT 1576
Qy 1518 AGAGATCTGAT 1577
Db 3387 AGAGATCTGAT 1636
Qy 1578 CAGCAG 1637
Db 3447 CAGCAG 1696
Qy 1638 TGGGTCTCCATCCAG 1697
Db 3507 TGGGTCTCCATCCAG 1756
Qy 1698 CATGAGATGAT 1757
Db 3567 AACTAGAGATGAT 1816
Qy 1758 CTGCGAG 1817
Db 3627 CTGCGAG 1876
Qy 1818 TTGCTTCAGAT 1859
Db 3687 CCTCTGCAAGCAATGCGGAT 3728

RESULT 12
AAQ35251
ID AAQ35251 standard; cDNA: 5406 BP.
AC AAQ35251:
XX
DT 25-JUN-1993 (first entry)
XX
DE Human flk-1 coding sequence.
XX

Human: receptor; protein; tyrosine kinase; pTK; primitive; mammalian;
hematopoietic cell; ptk; mature; mhc; fetal; liver kinase 2; flk-1;
liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;
multipotent; T-lymphoid; lineage; stomach; kidney; lung; heart;
intestine; muscle; lymph node; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FI CDS 208..4311
FI /*tag= a
PV KW9300349-A.
XX
PD 07-JAN-1993.
XX
PF 26-JUN-1992: 92WO-US05401.
XX
F8 28-JUN-1991: 91US-0728913.
PR 15-NOV-1991: 91US-0793065.
PR 24-DEC-1991: 91US-0813593.
PE 02-APR-1992: 92WO-US02750.
XX
PA (UWPR-) UNIV PRINCETON.
XX
PI Lemischka JB.
XX
XX WPI: 1998-036323/04.
DR P-FSDS: AAR1377.
XX
XX Nucleic acid encoding receptor protein tyrosine kinase - allows
development of ligands to stimulate proliferation and/or
differentiation of mammalian hematopoietic stem cells
XX
PS Claim 10: Fig 2: 78pp: English.
XX
CC This sequence encodes a human receptor protein tyrosine kinase which
belongs to a new functional class of protein tyrosine kinases (PTKs).
However, this PTK is not in the same class as flk-2. PTKs in the same
class as flk-2 are expressed in primitive mammalian hematopoietic
cells but not in mature hematopoietic cells (mhc). This gene
is expressed in more mature hematopoietic cells. The protein encoded
by this sequence is an example of a receptor PTK and is called fetal
liver kinase 1 (flk-1). flk-1 is expressed in fetal liver, spleen,
thymus, brain, stomach, kidney, lung, heart and intestine, and adult
brain, kidney, heart, spleen, lung, muscle, marrow and lymph nodes.
XX
SQ Sequence 5406 BP: 1412 A: 1299 C: 1422 G: 1273 T: 0 other;
Query Match 33.7%; Score 708.4; DB 14: Length 5406;
Seqs: Local Similarity 65.0%; Pred. No. 1.3e-204;
Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;
Qy 201 CACTCTTAATCTTACCAATCATGATTTCTCTGCAAGATTGAGGACCTATGCTGACAG 250
Db 2070 CATCTTGAATGTGCGATTTTCAAGATTCCTCTCTGAGGACCAAGGCGACTATGTTCTC 2129
Qy 261 AGCCAGGAGATGATACAGAGGAGGAGAAATCTCCAGAGAGAGAAATTAACAATCAGAGA 320
Db 2130 TGCATCAAGATAGAGAGACCAAGAAAGACATTCCTGGTCAACAGCTCATCTCTAGTA 2189
Qy 321 TACGAGAGACCATCTCTCTGCAAGACCTCAGTATCACAGATGGCCATCAGCAGCTTC 380
Db 2190 GGGCATGSCACCCATGATCAGCGGAAATCTGGAGATACAGACAAACCAATTCGGAGAC 2249
Qy 381 CACCACTTTAGATGTCATGCTTATGTTGTCCTCCGAGCCTCAGATCCTGTTTAAAAA 440
Db 2250 CATTAAGATGATGTCCTGAGATTCCTGGAATTCCTACCCACATATGATGTTCAAGA 2309
Qy 441 CAACACAGAT 500
Db 2310 CAACGAT 2369

QY 501 TATTGAAGAGTACAGAGAGAGATGAGGTGTCTATCATCTGCAAGAGCCACCAAGAGAA 560
 DB 2370 TATCCGAGGTGAGGAAGAGATGAGGTGTCTATCATCTGCAAGAGCCACCAAGAGAA 2429
 QY 561 GGGTCTGTGGAGAGTTCAGCAATACCTGCTGTCAGCAAGAGCCACCAAGAGAA 620
 DB 2430 TGGCTGTGCAAG 2489
 QY 621 GAGGTGATCATCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 580
 DB 2490 GGAAGTATTAATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2549
 QY 681 CTTCTTTATCCGAAG 737
 DB 2550 CATTCCTGAG 2609
 QY 738 AATTATTAATGAG 797
 DB 2610 TATGTCATGATCCAG 2669
 QY 798 TCCAG 857
 DB 2670 TCCAG 2729
 QY 858 GCTTTTGAAG 917
 DB 2730 TGCCTTGGCAG 2789
 QY 918 GACTGTGCTGTGAAATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 977
 DB 2790 AACAGTACGCTGAAGATGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2849
 QY 978 GACTGAGTAAATCTTGACCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1037
 DB 2850 GTCTGAATCTGAAGTCTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2909
 QY 1038 AGCTTGACAG 1097
 DB 2910 CGCTTGACAG 2969
 QY 1098 TCTTCCAGTACCTGCAAG 1157
 DB 2970 CTTATCAATCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3029
 QY 1158 ACAGATGAG 1217
 DB 3030 AGCTTGACAG 3086
 QY 1218 ACTAGATAGCTGACAG 1277
 DB 3087 CTGAG 3146
 QY 1278 TCTGAGTATGTTGAG 1337
 DB 3147 GCTGAGTATGTTGAG 3206
 QY 1338 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAG 1397
 DB 3207 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAG 3266
 QY 1398 AAGTGCATCTATGAG 1457
 DB 3267 GAAGTGTATCCAG 3326
 QY 1458 GAAGATTTGATTTTGGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1517
 DB 3327 TAAGATCTGATTTGAG 3386
 QY 1518 AGAGATCTGATTTGAG 1577
 DB 3387 AGAGATCTGATTTGAG 3446
 QY 1578 CAGCAG 1637

DB 3447 CACAATTCAG 3506
 QY 1638 TGGGTCTCCATACCCAGAGAGATACAAATGAGATGAGAGAGATTTTGCAGTTCGCTGAGAGAGAG 1697
 DB 3507 TGGCTTCCCATACCCCTGGGTCAAGATTTGATGAGAGATTTTGTAGAGATTTGAAGAAGG 3566
 QY 1698 CATGAGGATGAGAGAGTCTTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1757
 DB 3567 AACTAGATGAGAGAGTCTTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3626
 QY 1758 CTGCGACAG 1817
 DB 3627 CTGCGACAG 3686
 QY 1818 TTTGCTTCAAGCAATTTTCAAGCAATTTTCAAGCAATTTTCAAGCAATTTTCAAGCAATTTTCAAG 1879
 DB 3687 CCTCTGCAAGCAATTTTCAAGCAATTTTCAAGCAATTTTCAAGCAATTTTCAAGCAATTTTCAAG 3728
 RESULT 13
 ID AAG40916 standard: cDNA: 5406 BP.
 XX AAG40916
 AC AAG40916
 XX 12-OCT-1993 (first entry)
 DE Murine flk-1 cDNA.
 XX
 DE Murine; receptor; protein; tyrosine kinase; ptk; flk-1; flk-2; adult;
 KW family; primitive; hematopoietic cell; mature; fetal; liver; spleen;
 KW thymus; brain; stomach; kidney; lung; heart; intestine; bone marrow;
 KW muscle; lymph node; ss
 XX MUS musculus.
 XX
 XX Key Location/Qualifiers
 FH 208..4311
 FT CDS /tag- a
 FT sig_peptide 208..264
 FT /tag- b
 FT /note- "Hydrophobic leader sequence"
 FT mat_peptide 265..4308
 FT /tag- c
 XX W09310136-A.
 XX
 XX 27-MAY-1993.
 XX
 XX 16-NOV-1992: 92WO-US09893.
 XX
 XX 15-NOV-1991: 91US-0793065.
 XX (TYPR-) UNIV PRINCETON.
 XX
 XX Lewischka IR:
 XX
 XX NP1: 1993-182479/22.
 XX P-FSDB: AAB37504.
 XX
 XX Tctipotent haematopoietic stem cell receptors, their ligands and
 FT DNA sequences - for treating anaemia(s) and bone marrow damage
 FT due to e.g. cancer chemotherapy or radiotherapy
 XX
 XX Claim 14: Fig 2: 127pp: English.
 XX
 XX This sequence encodes the murine receptor protein tyrosine kinase
 CC (PTK), flk-1. This PTK is not in the same family as flk-2 (see also
 CC A240914-15) as it is expressed in primitive hematopoietic cells and
 CC also in mature hematopoietic cells. flk-1 is expressed in fetal
 CC liver, spleen, thymus, brain, stomach, kidney, lung, heart and
 CC intestine and adult brain, bone marrow, kidney, heart, spleen, lung,
 CC

CC muscle and lymph nodes.

CC muscle and lymph nodes.

XX Sequence 5406 BP: 1412 A: 1298 C: 1423 G: 1273 T: 0 other:

QQ Query Match 33.7%: Score 708.4: DB 14: Length 5406:
Best Local Similarity 55.0%: Pred. No. 1.3e-204:
Matches 1080: Conservative 0: Mismatches 576: Indels 6: Gaps 2:

QY	201	CACCTTTAACTTACCAATCAAGAAATGTTCCCTGCAAGATTCAGGACCACTATGCGTCGAC	260
DB	2070	CATCTTGATTTGGACATTCAGAAATGCCCTCTCTGAGGACCAAGCGCAATGTTGCTC	2129
QY	261	AGCCAGGAATGTATACACAGGCGAGAAATCCTCCAGAAAGAAATATACAAATCAGAGA	320
DB	2130	TGCTCAAGATAGAGACCAAGAAAGAGATGCTCGGTCAAAAGCTCATCACTTAGA	2189
QY	321	TCAGGAGCAACATACCTCTCTGGAAGCTCAGTGATCACACAGTGGCCATCAGCAGTTC	380
DB	2190	GGCATGCGCAATGATATACCGGAAATTCGGAAATCAGACAAACCATTTGCGGAGAC	2249
QY	381	CACCACCTTAGACTGTACGTAAATGGTGTCCCGAGCCTCAGATCAGTCTTGAATAA	440
DB	2250	CATGAGTGACTTGCCAGCATCTGGAAATCTTACCCACACATTAAGTGTTCAAAGA	2309
QY	441	CAACCAAAAAATACAAAGAGCCTGGAAATTAATTTAGACACAGGAGCAACAGCTGTT	500
DB	2310	CAACGAGACCTGTGAGAGATCAGCAATGTACTGAGAGATGGGACCGGACCTTGAC	2369
QY	501	TATGAAAGAGTCCACAGAGAGATGAGGTGTCTATCTGCAAGGCCACACACAGGAA	560
DB	2370	TATCCGAGGTGAGGAGAGATGAGGAGCTCTACACCTCCAGGCTCGCAATGTCCT	2429
QY	561	GGGCTCTGTGAAAGTTCAGCATACCTCACTGTTCAGGAACTCGGACAGCTGATCT	620
DB	2430	TGGCTGTCAAGAGCGGAGACGCTCTTCAATAGAGGTGCCCAGGAAAGACCACTT	2489
QY	621	GGAGCTGATCACTTAACATCAGCTGTGCGGTGCGACTCTCTCTGCTCTCTATTAAC	680
DB	2490	GGAGTCAATATCTGTGCGCACTGCAAGTGAATGCTTCTCTGCTCTCTTGT	2549
QY	681	CTCTCTTAACGAAAAATGAAAGGTC---TCTCTTGAAATAGAGATGACTATC	737
DB	2550	CAATCTGTAGCGACCGTTAAGCGGCCAATGAGGGGAACTCAAGACAGCTACTTGT	2609
QY	738	AAATTAATAGCCAGATGAGTCTCTTTGATGACAGGTGTAGCGGTCTCCCTTAAGA	797
DB	2610	TATGTGATGATCCAGATGAATTTGCCCTTGGATGAGCGCTGTGAACGCTTGCTTA	2669
QY	798	TCCCAAGAGTGGAGTGTTCGGGAGAGACTTAACTGGGCAATCAGTGGAGAG	857
DB	2670	TCCAGCAAGTGGAAATCCCCAGGACCGGCTGAACATGAGAACTCTTGGCGCGG	2729
QY	858	GGCTTTTGGAAAGTGTTCAGCATCAGCATTTGGCATTAAGAAATCAGTGGCG	917
DB	2730	TGCCCTTCGGCAAGTGAATGAGGAGAGCGTGTGGAATGACAGACAGCGACTTGA	2789
QY	918	GACTGTGCTGTGAAATGCTGAAGAGGGGCGCACGGCGAGGTACAAAGCTGTAT	977
DB	2790	AACAGTAGCGGTGAGATGTGAAGAGAGAGACACACAGGAGATCGAGCCCTCAT	2849
QY	978	GACTGAGCTTAAATCTTACCCACATTTGCCACCACTCTGAACTGGTTAACTGTGG	1037
DB	2850	GTCTGAGCTCAAGTCTCATCCCATTTGCTACCAATGTCAATGTGTGAACCTCTAG	2909
QY	1038	AGCCTGACCAAGCAGAGAGGCTCTGTATGTTGATTTGATAGTATGAAATAGGAA	1097
DB	2910	CGCTGCGACCAAGCGGAGGCGCTCATGTGATGTGGAAATTCGAAATTTGGAA	2969
QY	1098	TCTCTCACTACCTCAAGACCAAGCTGACTTAATTTTCTCAACAGGATGCACT	1157
DB	2970	CTATCACTTACTTACGGGCAAGAGATGATTTGTTCCTATAGAGCAAGGGCG	3029
QY	1158	ACATGAGGCTTAAAGAAATGAGGACGAGCTGGACAAAGCGCAAGAACCAAG	1217

Db 3030 AGCTTCGCGCAGGCAAGGACTAGCTTGGG---GAGCTCTCCGTGGATCTGAAACG 3088

QY 1218 ACATAGATAGGCTCACAGCAGCAAGCTTTGCGAGCTCCGCTTCAGGAGATTAAG 1277

Db 3087 CTGAGCAGCAACACAGCAGCGAGCTTGGCAGCTCAGGCTTTGTTGAGGAGAAATC 3146

QY 1278 TCTGAGTGTGTTGAGGAGAGAGGATTTCTGACGGTTCCTACAGAGGCCATCACTAT 1337

Db 3147 GCTCAGTGTATGAGGAGAGAGAGAGTCTCTGAAGAACTGTACAGGACTTCTGACCT 3206

QY 1338 GGAAGATCTCAATTTCTTACAGTTTCAAGTGGCCAGGCGATGGAGTTCCTGTCTCCAG 1397

Db 3207 SGAGCATCTCACTGTTACAGCTTCCAGTGGCTAAGGCTGAGGTTCCTTGGCATCAAG 3266

QY 1398 AAGTGCATTCATCGGGAGCTGGCAGCGAGAAACATTTCTTTCTGAGAAACAGCTGT 1457

Db 3267 GAAGTGTATCCAGCGGACTTGGCAGCAACAAACATCTCTATCGGAGAGAGTGTGT 3326

QY 1458 CAAGATTTGATTTTGGCTTGGCCGGGATATTTATAGAACCCCGATTTATGTGAGAA 1517

Db 3327 TAAATCTGTGACTTGGCTTGGCCGGGACATTTATAGAACCCCGATTTATGTGAGAA 3386

QY 1518 AGGAGTACTTCGACTTCTCTGAAATGGATGGCTCCCGAATCTATCTTTACAAAATCTA 1577

Db 3387 AGGAGATGCCGACTCCCTTTGAAGTGGATGGCCGCGAAACCAITTTTGCACAGTATA 3446

QY 1578 CAGCACCAGCAGCGAGCTGTGCTTACGGAGTATGCTGTGGGAAATCTTCCTTAGG 1637

Db 3447 CACAAITCAGCGAGTGTGCTTTCGCTGTGCTGTGGGAAATATTTTCCTTAGG 3506

QY 1638 TGGCTTCCATCCAGGAGTACAAATGAGTGGAGCTTTTGCAGTCCGCTGAGGAGG 1697

Db 3507 TGGCTTCCATCCGCTTGGGCTCAAGTGTGAAGAAATTTTGTAGGAGTTGAGAGAG 3566

QY 1698 CATCAGATGAGAGCTTCGAGTACTCTACTCTCTGAAATCTATCATCATCATCATGCT 1757

Db 3567 AACTAGATCCGGCTTCTGACTACTACCTCCCGGAAATGTACACAGGATCTGAGCTG 3626

QY 1758 CTGGCAGAGACCCAAAGAGGCGCAAGATTTGCAAGCTTTGTGAAAACTAGGTGA 1817

Db 3627 CTGGCATGAGGACCCCAACAGAGACCTCTCTTTCAGGTGTGGAGCTTGGGAA 3686

QY 1818 TTTGCTTCAAGCAATGTATACACAGAGTGTAAAGCTACAT 1859

Db 3687 CTCTTGCAGCAATGCGCAGGAGTGTGCAAGACTATAT 3728

RESULT 14

AAQ81014

ID AAQ81014 standard: cDNA; 5406 BP.

XX AC AAQ81014;

XX DT 18-AUG-1995 (first entry)

XX DE Fkl1 receptor protein-tyrosine-kinase cdna.

XX KW Mouse Fkl1; receptor protein-tyrosine-kinase; hematopoietic stem cell; fetal liver kinase; ds.

XX CS Mus musculus.

XX FH Key Location/Qualifiers

XX CDS 208..4311

XX FI /*tag= a

XX FI /product= fkl1 receptor protein-tyrosine-kinase

XX FI 208..264

XX FI /*tag= b

XX FI mat_peptide 265..4308

XX FI /*tag= c

XX FN R03500554-A

XX 05-JAN-1995.
 XX 17-JUN-1994; 94HO-US06944.
 XX 18-JUN-1993; 93US-0080244.
 PR 21-JUN-1993; 93US-0081508.
 PR 23-NOV-1993; 93US-0157490.
 XX (UTPR-) UNIV PRINCETON.
 XX Lemischka IR;
 PI WPI; 1995-052014/07.
 DR P-PSDB; AAR67817.
 XX
 PT Ligand for receptor protein tyrosine kinase - useful for the
 PT stimulation of primitive hematopoietic stem cells causing
 PT proliferation and/or differentiation
 XX
 CC The sequence corresponds to a cDNA encoding a mouse Flk1 (fetal
 CC liver kinase) receptor protein-tyrosine-kinase. Flk1 is expressed
 CC in both primitive and mature hematopoietic cells, and in other
 CC tissues. The gene product is useful in isolation of receptor
 CC ligands, which have applications in diagnosis of bone marrow
 CC disorders and in stimulating proliferation and/or differentiation
 CC of primitive hematopoietic stem cells.
 XX
 SQ Sequence 5406 BP; 1412 A; 1298 C; 1423 G; 1273 T; 0 other;
 Query Match 33.7%; Score 708.4; DB 16; Length 5406;
 Best Local Similarity 65.0%; Pred. No. 1.3e-204;
 Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;
 OY 201 CACTCTTAATCTTACCATCAATGATTTTCCCTGCAAGATTCAGGCACCTATGCGTCGAC 260
 DB 2070 CATCTGTATGTTGGCAATTCAGATGCGCTCTCTGAGGACCAAGCGGCTATGTTGTC 2129
 OY 261 AGCCAGGAATGTATACACAGGGGAGAAATCTCCAGAGAAAGAAATATACATCAGAGA 320
 DB 2130 TGCTCAAGATAGAGACCAAGAAAGACATTCCTGTTCAAGAGCTCATATCTAGA 2189
 OY 321 TCAGGAGACCATACCTCTGCGAAACCTCAGTATCACACAGTGGCCATCAGAGTTC 380
 DB 2190 GCGATGGCCACCAATGATCCCGGAAATCTGGAGATCAGACAAACCAATGGCGAGC 2249
 381 CACCACCTTAGACTGTATCTATGCTGATGCTGCTCCCGAGCTCAGATCTGTTTAAAAA 440
 2250 CATTGAAGTGACTTGGCCAGCATCTGGAAATCTTACCCACACATACATGTTCAAGA 2309
 OY 441 CACACAAATATACACAGAGCTGGAAATTTATTTAGGACAGGAGCAGCAGCTGT 500
 DB 2310 CAACGAGACCTTGGTAGAAGATTCAGGATTTACTGAGAGATGGGAACCGGACCTGAC 2369
 OY 501 TATTGAAGAGTTCACAGAGAGATGAAGTGTCTATCAGTGCAGCTCAGATCTGTTTAA 560
 DB 2370 TATCCGAGGGTGGAGAGGATGAGGCTCTTACACCTTGGCCAGGCTGCAATGTCT 2429
 OY 561 GGCTCTGTGGAAATTCAGCATACCTGCTTCAAGGAACCTTGGCAAGATCTAATCT 620
 DB 2430 TGCTGTGCAAGAGGGAGAGCTCTTCAATATAGAGTGGCCGGAAGAACCACTT 2489
 OY 621 GGAGCTGATCATTACATGACCTGTGTGGCTGGACATCTCTTGTGCTCTTAAAC 680
 DB 2490 GGAAGTCAATATCTCTGCTGGGCACTCAGTATGATGCAATGTCTTGTGCTCTTGT 2549
 OY 681 CTTCTTATCCGAAATGAAGGTC---TCTCTGAAATATACACTGATCTACTATC 737
 DB 2550 CATTCTGTAGGACCGTTAGCGGGCCCAATGAAGGGGAATGAAGACAGGCTACTT 2609
 OY 738 AATTATATGACCCAGATGAAGTTCCTTTGGATGAGCAGTGTGAGCGCTCCCTTATGA 797

DB 2610 TATTGTATGATCCAGATGAATGGCCCTTGGATGAGCGCTGTGACGCTTGCCTTATGA 2669
 OY 738 TGCAGCAAGTGGAGTITGCCGGGAGAGACTTAACTGGGCAAACTACTTGAAGAG 857
 DB 2570 TGCAGCAAGTGGAAATTCGCCAGGACCGGCTGAAACTAGGAAACCTCTTGGCCGCG 2729
 OY 858 GGCCTTTGGAAATGGTTCAGCATCAGCATTTGGCATTTAGAAATCACCCTACCTGCGCG 917
 DB 2730 TGCCTTGGCCAGTATGAGGAGAGGCTTTTGGATTTGACAGACGAGCTTGC 2789
 OY 918 GACTGTGCTGTGAAATGCTTGAAGAGGGGCGCCAGCGAGTACAAAGCTCTGAT 977
 DB 2790 AACAGTAGCCGTCAAGATGTTGAAGAAGAGCAACACACAGCGAGCATGAGCCCTCAT 2849
 OY 978 GACTGAGCTAAATATCTTACCCACATTTGCCACCATCTGACGTGTTAACTTCTGGG 1037
 DB 2850 GTCTGAGACTCAAGATCTTCCACATTTGGTCAATCTCAATGCTGTAACCTCTAGG 2909
 OY 1038 AGCTTGACCAAGAGGAGGCGCTCTGATGTTGATGTTGTAATCTGCAAAATATGAAA 1097
 DB 2910 CGCTGACCAAGCGGGGCGCTCTCATGTTGATTTGGNATTTCTGAGTTTGGAAA 2969
 OY 1098 TCTCTCAACTACCTCAAGACAACTGACTTATTTTCTCAACAGGATCGACACT 1157
 DB 2970 CCTATCACTTACTTACGGGCAAGAGAAATGAAATTTGTTCCCTATAGAGCAAGGGC 3029
 OY 1158 ACATGAGGCGCTTAAGAAGAAATATGAGCGGCTGGACAAAGCAAGCAAGCAAG 1217
 DB 3030 AGCTTCCGCCAGGCAAGGACTACGTTGG---GAGCTCTCCGTGGATCTGAAAAGCG 3086
 OY 1218 ACTAGATAGCTTACACAGCGGAAAGCTTTTGGAGCTTCCGGCTTTTTCAGGAATAAAG 1277
 DB 3087 CTGGAGACATCATCCAGCAGCCAGAGACTCTGCCAGCTCAGGCTTTTGTGAGGAAATC 3146
 OY 1278 TCTGAGTATGTTGAGGAGAGGAGGATCTGACGCTTCTACAGAGGAGCCCTACATAT 1337
 DB 3147 GCTCAGTATGAGAGAGAGAGAGCTTCTGAAGAACTGTACAGGAGCTTCTGACCTT 3206
 OY 1338 GGAGATCTGATTTCTTACAGTTTTCAGTGGCCAGAGGAGCTGAGTCTCTGCTTCCAG 1397
 DB 3207 GGAGCACTCATCTGTTACAGCTTCCAAAGTGGCTAAGGCGATGGAGTTCTTGGCATCAAG 3266
 OY 1398 AAGTGCATTCATCGGACCTTGGCAGCGAGAAACATTTCTTTATCTTGAGAAACAGTGT 1457
 DB 3267 GAAGTATTCACAGAGGAGCTTGGCAGCAGCAACATCTCTCTATCGSAGAGAAATGTGT 3326
 OY 1458 GAAGATTTGTGATTTTGGCTTGGCCGGGATATTTAAGAAACCCGATTTATGTGAGAAA 1517
 DB 3327 TAAGATCTGTGACTTGGCTTGGCCGGGACATTTAAGACCCCGGATTTATGTGAGAAA 3386
 OY 1518 AGGAGATCTGACTTCTCTGAATGGAATGGCTCCCGATCTATCTTTGACAAAATCTA 1577
 DB 3387 AGGAGATCGGAGCTCTCTTGAAGTGGATGGCCCGGAAACCACTTTTTCAGAGATATA 1637
 OY 1578 CAGCAGAGAGGAGCTGTGGCTTACCGAGTATTTGCTGTGGAAATCTTCTCTTAGG 1637
 DB 3447 CACAATTCAGAGGAGTGTGGCTTTCGGTGTGTGCTGTGGAAATATTTTCTTAGG 3506
 OY 1638 TGGGTCTCATACCGAGGATCAAAATGGATGAGGACTTTTTCAGTGGCTGAGGGAAG 1697
 DB 3507 TGGCTTCCCATACCTTGGGCTAAGATTTGATGAAGAAATTTTGTAGAGATTCAGGAAGG 3566
 OY 1598 CATGAGATGAGAGCTCTGAGTACTCTTACTCTGAATCTATCAGATCATCTGCTGACTG 1757
 DB 3567 ACTAGATGCGGCTCTGACTACTACCTCCCAAGAAATGTACAGACCATCTGCTGACTG 3626
 OY 1758 CTGGCAGAGAGACCAAGAAAGAGGCGGAGATTTGCAGAACTTGTGGAATAACTAGTGA 1817
 DB 3627 CTGGCATGAGAGCCCAACAGAGAGCTCTGTTTTCAGAGTGTGTGGAGCATTTGGGAA 3686
 OY 1818 TTTGCTTCAACCAATGTACACAGGATGTTAAGACTACAT 1859

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 09:21:42 ; Search time 2723.29 seconds
(without alignments)
12494.717 Million cell updates/sec

Title: X51602_COPY_1900_4000

Perfect score: 2101

Sequence: 1 TATATACAGATGTGCCAAA.....CGACAGCAGCAGCAGCIGIGG 2101

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:**

1: em_estb:**
2: em_estb:**
3: em_estb:**
4: em_estb:**
5: em_estb:**
6: em_estb:**
7: em_estb:**
8: em_estb:**
9: gb_est1:**
10: gb_est2:**
11: gb_est3:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: gb_gss:**
18: em_gss_hum:**
19: em_gss_inv:**
20: em_gss_pin:**
21: em_gss_vrt:**
22: em_gss_fun:**
23: em_gss_mam:**
24: em_gss_mus:**
25: em_gss_other:**
26: em_gss_pro:**
27: em_gss_fod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	827.6	39.4	3976	11 AK004721	Mus muscu
2	715	34.0	936	13 B1762725	B1762725 603048490
3	571.4	27.2	697	14 BQ604389	BQ604389 MI-P-Cp1-
4	561.4	26.7	666	9 AL042601	AL042601 DKF2p434L
5	558.2	26.6	650	12 BG005458	BG005458 MR3-GN018
6	466.8	22.2	1000	9 AL575294	AL575294 AL575294

7	451.6	21.5	566	14 BQ770907	BQ770907 UI-M-F10-
8	444.6	21.2	1005	9 AL570084	AL570084 AL570084
9	442.4	21.1	940	9 AL569839	AL569839 AL569839
10	423.4	20.2	855	9 A1326556	A1326556 mm09f03.y
11	415.2	19.8	512	12 BG384715	BG384715 303951.MA
12	406	19.3	606	13 B1458691	B1458691 603199829
13	385.6	18.4	561	9 A4066758	A4066758 mm09f03.r
14	377.4	18.0	527	9 AL571839	AL571839 AL571839
15	374.8	17.8	835	13 B1731060	B1731060 603351683
16	370.2	17.6	441	12 BG384720	BG384720 303961.MA
17	349.2	16.6	387	12 BF086406	BF086406 RC6-GN007
18	336.6	16.0	383	14 BQ387353	BQ387353 PM4-GN030
19	332.4	15.8	824	12 BG435852	BG435852 602508476
20	321.6	15.3	500	14 R62703	R62703 Y110B03.r1
21	321.2	15.3	645	10 AW167792	AW167792 XG56D06.x
22	321.2	15.3	652	9 A1800102	A1800102 T120F07.x
23	321.2	15.3	782	9 AL573250	AL573250 AL573250
24	321.2	15.3	797	12 BG571885	BG571885 602593233
25	320.8	15.3	955	9 AL576447	AL576447 AL576447
26	312.8	15.0	689	13 B3075253	B3075253 B3075253
27	312.2	14.9	656	10 B8577273	B8577273 B8577273
28	308	14.7	859	9 AL573279	AL573279 AL573279
29	306	14.6	581	9 A194109	A194109 Z138407.S
30	303	14.4	496	10 AW129059	AW129059 XE94f10.x
31	295.8	14.1	561	10 AW029025	AW029025 WY98f10.x
32	291	13.9	623	12 BF087145	BF087145 PM1-HT045
33	288.2	13.7	711	9 A880212	A880212 VX39D08.r
34	280.6	13.4	819	9 AL570470	AL570470 AL570470
35	278.8	13.3	672	9 A1761753	A1761753 W67H06.x
36	277.6	13.2	304	12 BG015691	BG015691 CM1-GN036
37	277.6	13.2	1067	14 BM925760	BM925760 AGENCOURT
38	274.2	13.1	681	14 W26612	W26612 36d4 Human
39	267	12.7	463	10 AW681286	AW681286 EST00004
40	266	12.7	671	9 A1568689	A1568689 AL568689
41	265.2	12.6	504	9 A1089100	A1089100 OM82q10.S
42	264.6	12.6	472	12 BF061039	BF061039 7162h11.x
43	260.4	12.4	547	10 BE030679	BE030679 128743.MA
44	259.6	12.4	320	9 A1716140	A1716140 UI-R-Y0-a
45	257.6	12.3	359	12 BF543282	BF543282 UI-R-Y0-a

ALIGNMENTS

RESULT 1
AK004721
LOCUS
DEFINITION
Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200012B22:FMS-like tyrosine kinase 1, full insert sequence.
ACCESSION
AK004721
VERSION
AK004721.2 GI:15919895
KEYWORDS
HIC: CAP trapper.
SOURCE
Mus musculus (strain: C57BL/6J) adult male lung cDNA to mRNA, clone:1200012B22.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
1
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
93279253
PUBMED
10345636

REFERENCE
2
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Suganara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159

PD XX 15-JUN-1999.
 PF XX 10-FEB-1998; 98US-0021324.
 PR XX 19-NOV-1992; 92US-0977451.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.
 PR 26-JUN-1992; 92US-0906397.
 PR 12-NOV-1992; 92US-0975049.
 PR 30-APR-1993; 93US-0055269.
 PR 31-OCT-1994; 94US-0252498.
 PR 15-FEB-1996; 96US-0601891.

(UYPR-) UNIV PRINCETON.
 Lemischka IR;
 MPI; 1999-357194/30.
 P-PSDB; AAY08618.

Isolating hematopoietic cells expressing fetal liver kinase 1
 receptors

Disclosure; Fig 2; 59pp; English.

This invention describes a novel method of isolating cells expressing fetal liver kinase 1 (flk-1) receptors on their surface and comprises the following steps: (a) providing a polyclonal or monoclonal antibody specific to the flk-1 receptor and isolating the cells that have bound to the antibody. The method can be used to isolate hematopoietic stem cells in any mammal but preferably a rat, mouse, rabbit or human. The proteins of the invention belong to the receptor protein family. This sequence encodes the murine flk-1 protein which is used in the method of the invention.

SQ Sequence 5404 BP; 1411 A; 1297 C; 1423 G; 1273 T; 0 other;

Query Match 24.8%; Score 74.6; DB 20; Length 5404;
 Best Local Similarity 78.8%; Pred. No. 1.8e-13;
 Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTTCCTGTAATGGATGGCTCTGTAATCTATCTTTGACAAAATCTACAGCACCA 128
 DB 3395 CCCGACTCCCTTTGAAGTGGATGGCCCGGAAACCATTTTTCACAGATATACACATTC 3454
 129 AGAGCGACGTGTGGTCTTACGAGTATCTGTTGGGAATCTTCCTTAGGT 181
 3455 AGAGCGATGTGGTCTTTCGGTGTGTTGCTCTGGGAATATTTTCCTTAGGT 3507

RESULT 12

AAQ29957
 ID AAQ29957 standard; cDNA; 5406 BP.

XX AC AAQ29957;

DT 15-MAR-1993 (first entry)

DE flk-1 cDNA sequence.

Thymidine kinase; TK; hematopoietic; stem cells; proliferation;
 differentiation; progenitor cells; foetal liver kinase; ss.

OS Mus musculus.

XX Key Location/Qualifiers
 FH 208...4311
 FT /*tag= a

PN W09217486-A.

XX

PD XX 15-OCT-1992.
 PF XX 02-APR-1992; 92KO-US02750.
 PR 02-APR-1991; 91US-0679666.
 PR 28-JUN-1991; 91US-0728913.
 PR 15-NOV-1991; 91US-0793065.
 PR 24-DEC-1991; 91US-0813593.

(UYPR-) UNIV PRINCETON.

Lemischka IR;

WPI; 1992-366185/44.

P-PSDB; AAR28041.

Stimulating proliferation and/or differentiation of primitive
 mammalian hematopoietic stem cells - using ligand that binds
 thymidine kinase and flk-1 and flk-2

Claim 14; Fig 2; 94pp; English.

The murine foetal liver kinase (flk-1) clone was isolated by
 standard PCR techniques from stem-cell receptor-contg. tissue cDNA
 libraries. Suitable tissues include foetal liver, spleen or thymus
 cells or adult marrow or brain cells. The PCR primers used are based
 on known sections of the flk-1 gene. The murine flk-1 clone may be
 used in a vector to transform hematopoietic cells. The thymidine
 kinase encoded by flk-1 is expressed in primitive but not mature
 hematopoietic cells. Ligand binding to the TK may be prepd. which
 can stimulate proliferation and/or differentiation of primitive
 hematopoietic cells in vivo. The ligands can stimulate the
 proliferation of additional primitive stem cells. Differentiation into
 more mature progenitor cells, or both.

See also AAQ29954-6.

SQ Sequence 5406 BP; 1412 A; 1294 C; 1427 G; 1273 T; 0 other;

Query Match 24.8%; Score 74.6; DB 13; Length 5406;
 Best Local Similarity 78.8%; Pred. No. 1.8e-13;
 Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTTCCTGTAATGGATGGCTCTGTAATCTATCTTTGACAAAATCTACAGCACCA 128
 DB 3395 CCCGACTCCCTTTGAAGTGGATGGCCCGGAAACCATTTTTCACAGATATACACATTC 3454
 129 AGAGCGACGTGTGGTCTTACGAGTATCTGTTGGGAATCTTCCTTAGGT 181
 3455 AGAGCGATGTGGTCTTTCGGTGTGTTGCTCTGGGAATATTTTCCTTAGGT 3507

RESULT 13

AAQ53504
 ID AAQ53504 standard; cDNA; 5406 BP.

XX AC AAQ53504;

DT 27-JUN-1994 (first entry)

DE Murine flk-1 cDNA.

Receptor protein tyrosine kinase; ptk family; foetal liver kinase;
 flk; primitive; totipotent; hematopoietic cell; stem cell;
 proliferation; stromal cell; ds.

OS Mus musculus.

XX Key Location/Qualifiers
 FH 208...4311
 FT /*tag= a

FT sig_peptide

FT /*product= flk-1
 FT /*tag= b

FT mat_peptide /note= "hydrophobic leader"
 FT 265..4308
 FT /*tag= c

XX US5270458-A.

XX 14-DEC-1993.

XX 02-APR-1991; 91US-0679666.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 26-JUN-1992; 92US-0906397.

XX 12-NOV-1992; 92US-0975049.

XX 19-NOV-1992; 92US-0977451.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1993-405021/50.

XX P-PSDB; AAR44996.

XX Isolated nucleic acid molecules of hematopoietic stem cell

XX receptor flk-2 - encoding mammalian receptor protein tyrosine

XX kinases expressed in primitive haematopoietic cells

XX Disclosure; Fig 2; 60pp; English.

XX Nucleic acid sequences coding for murine flk-2 and specified

XX subfragments of it are claimed. The murine flk-1 coding sequence

XX (i.e. A053504) is also disclosed. The flk polypeptides are

XX receptor protein tyrosine kinases which are expressed only in

XX primitive haematopoietic cells.

XX Sequence 5406 BP; 1412 A; 1298 C; 1423 G; 1273 T; 0 other;

XX Query Match 24.8%; Score 74.6; DB 14; Length 5406;

XX Best Local Similarity 78.8%; Pred. No. 1.8e-13;

XX Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

XX QY 69 CTCGACTTCCTTGAAGTGGTCTCTGAAATCTATCTTTCGACAAATCTACAGCACCA 128

XX Db 3395 CCCGACTCCCTTTGAAGTGGTGGCCCGGAAACCAATTTTGACAGATATACAAATTC 3454

XX QY 129 AGACGACGTGGTCTTACGAGATATGCTGTGGGAAATCTTCCTTAGGT 181

XX 3455 AGACGATGTGTGCTTTCGCTGTGCTGTGGGAAATATTTTCCTTAGGT 3507

XX RESULT 14

XX AAQ35251

XX ID AAQ35251 standard; cDNA; 5406 BP.

XX AC AAQ35251;

XX 25-JUN-1993 (first entry)

XX Human flk-1 coding sequence.

XX Human; receptor; protein; tyrosine kinase; pTK; primitive; mammalian;

XX hematopoietic cell; pHC; mature; mHC; fetal; liver kinase 2; flk-1;

XX liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;

XX multipotential; T-lymphoid; lineage; stomach; kidney; lung; heart;

XX intestines; muscle; lymph node; ss.

XX PN WC9300349-A.

XX 07-JAN-1993.

XX 26-JUN-1992; 92WC-US05401.

XX 26-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 02-APR-1992; 92WC-US02750.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1993-036323/04.

XX P-PSDB; AAR31377.

XX Nucleic acid encoding receptor protein tyrosine kinase - allows

XX development of ligands to stimulate proliferation and/or

XX differentiation of mammalian haematopoietic stem cells

XX Claim 10; Fig 2; 78pp; English.

XX This sequence encodes a human receptor protein tyrosine kinase which

XX belongs to a new functional class of protein tyrosine kinases (pTKs).

XX However, this pTK is not in the same class as flk-2. pTKs in the same

XX class as flk-2 are expressed in primitive mammalian hematopoietic

XX (pHC) cells but not in mature hematopoietic cells (mHC). This gene

XX is expressed in more mature hematopoietic cells. The protein encoded

XX by this sequence is an example of a receptor pTK and is called fetal

XX liver kinase 1 (flk-1). flk-1 is expressed in fetal liver/spleen,

XX thymus, brain, stomach, kidney, lung, heart and intestine, and adult

XX brain, kidney, heart, spleen, lung, muscle, marrow and lymph nodes.

XX Sequence 5406 BP; 1412 A; 1299 C; 1422 G; 1273 T; 0 other;

XX Query Match 24.8%; Score 74.6; DB 14; Length 5406;

XX Best Local Similarity 78.8%; Pred. No. 1.8e-13;

XX Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

XX QY 69 CTCGACTTCCTTGAAGTGGTCTCTGAAATCTATCTTTCGACAAATCTACAGCACCA 128

XX Db 3395 CCCGACTCCCTTTGAAGTGGTGGCCCGGAAACCAATTTTGACAGATATACAAATTC 3454

XX QY 129 AGACGACGTGGTCTTACGAGATATGCTGTGGGAAATCTTCCTTAGGT 181

XX 3455 AGACGATGTGTGCTTTCGCTGTGCTGTGGGAAATATTTTCCTTAGGT 3507

XX RESULT 15

XX AAQ40916

XX ID AAQ40916 standard; cDNA; 5406 BP.

XX AC AAQ40916;

XX 19-OCT-1993 (first entry)

XX Marine flk-1 cDNA.

XX Marine; receptor; protein; tyrosine kinase; pTK; flk-1; flk-2; adult;

XX family; primitive; hematopoietic cell; mature; fetal; liver; spleen;

XX thymus; brain; stomach; kidney; lung; heart; intestine; bone marrow;

XX muscle; lymph node; ss.

XX Mus musculus.

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

```
FT      mat_peptide      /note= "Hydrophobic leader sequence"
FT      265..4308
FT      /*tag= C
XX
XX      PN      WO9310136-A.
XX      XX
XX      PD      27-MAY-1993.
XX      XX
XX      PF      16-NOV-1992; 92WO-US09893.
XX      XX
XX      PR      15-NOV-1991; 91US-0793065.
XX      XX
XX      PA      (UYPR-) UNIV PRINCETON.
XX      XX
XX      PI      Lemischka IR;
XX      XX
XX      DR      WPI: 1993:182479/22.
XX      DR      P-PSDB; AAR37504.
XX      XX

Totipotent haematopoietic stem cell receptors, their ligands and
DNA sequences - for treating anaemia(s) and bone marrow damage
due to e.g. cancer chemotherapy or radiotherapy

Claim 14; Fig 2; 127pp; English.

This sequence encodes the murine receptor protein tyrosine kinase
(pTK), flk-1. This pTK is not in the same family as flk-2 (see also
AAQ40914-15) as it is expressed in primitive hematopoietic cells and
also in mature hematopoietic cells. flk-1 is expressed in fetal
liver, spleen, thymus, brain, stomach, kidney, lung, heart and
intestine and adult brain, bone marrow, kidney, heart, spleen, lung,
muscle and lymph nodes.

SQ      Sequence 5406 BP; 1412 A; 1298 C; 1423 G; 1273 T; 0 other;

Query Match      24.8%; Score 74.6; DB 14; Length 5406;
Best Local Similarity 78.8%; Pred. No. 1.8e-13;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Caps 0;

QY      69      CTCGACTTCCTGTAATGATGGCTCCTGATCTATCTTTGACAAATCTACAGACCA 128
          ||||| ||| ||||| ||||| ||| ||| ||||| ||| |||||
Db      3395      CCGGACTCCCTTGAAGTGGATGGCCCGGAAACCATTTTGGACAGTATACAAATTC 3454

QY      129      AGAGCGACCTGCTCTTACGAGATGCTGTGGGAAATCTTCCTTAGGT 181
          ||||| ||||| ||||| ||| ||||| ||||| ||| |||||
Db      3455      AGAGCGATGTGCTCTTCGGTGTGTGCTCTGGGAAATATTTCCITAGGT 3507

Search completed: December 8, 2002, 06:01:53
Job time : 360.95 secs
```


GenCore version 5.1.3
Copyright (C) 1993 - 2002 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: December 8, 2002, 05:16:46 ; Search time 2350.03 Seconds
(without alignments)
3665.203 Million cell updates/sec

Title: US-09-778-900a-3_COPY_200_500
Perfect score: 301
Sequence: 1 ggacactcgggtgaatg.....taacaaaaactcttaagt 301

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pi:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pi:*

25: em_ro:*

26: em_sts:*

27: em_un:*

28: em_vi:*

29: em_hgt_hum:*

30: em_hgt_inv:*

31: em_hgt_inv:*

32: em_hgt_other:*

33: em_hgt_mus:*

34: em_hgt_pln:*

35: em_hgt_rtd:*

36: em_hgt_nam:*

37: em_hgt_vrt:*

38: em_sy:*

39: em_hgt_hum:*

40: em_hgt_mus:*

41: em_hgt_other:*

number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	299.4	99.5	164519	9	AL138712 Human DNA
C 2	124.2	41.3	101987	2	AC128808 Rattus no
C 3	124.2	41.3	144090	2	AC128409 Rattus no
C 4	114.4	38.0	4017	6	AX481481 Sequence
C 5	114.4	38.0	4017	9	AF063657
C 6	112.8	37.5	7680	9	HSFLT
C 7	104.4	34.7	220965	2	HSAL312688
C 8	102.8	34.2	197279	9	AC104439
C 9	95.8	31.8	4734	10	RATFLT1
C 10	95.8	31.8	5482	10	MUSRTKB
C 11	95.8	31.8	6055	10	MUSRTKB
C 12	95.8	31.8	6275	10	D88689
C 13	94.2	31.3	666	4	AF233077
C 14	88	29.2	4272	5	AB065372
C 15	80.6	26.8	211330	10	AL646088
C 16	77.2	25.6	459	4	BIVGFPREC
C 17	76.2	25.3	4044	6	AX481480
C 18	76.2	25.3	4071	6	AR142803
C 19	76.2	25.3	4071	6	AR201382
C 20	76.2	25.3	4071	6	AX481483
C 21	76.2	25.3	4225	6	AX133539
C 22	76.2	25.3	4230	9	HSGFRK
C 23	76.2	25.3	4230	9	HUMKDRZ
C 24	76.2	25.3	4236	6	AR030758
C 25	76.2	25.3	5830	9	AF035121
C 26	75.4	25.0	163017	2	AC131433
C 27	74.6	24.8	4790	10	S53103
C 28	74.6	24.8	5391	10	MMFLK1
C 29	74.6	24.8	5406	6	AR005213
C 30	74.6	24.8	5406	6	AR071705
C 31	74.6	24.8	5406	6	AR071705
C 32	74.6	24.8	5406	6	I25171
C 33	74.6	24.8	5406	6	I40602
C 34	74.6	24.8	5464	10	BC020530
C 35	74.6	24.8	5470	6	AR068047
C 36	74.6	24.8	5470	6	AR125598
C 37	74.6	24.8	5470	10	MMFLK1M
C 38	74.2	24.7	73760	2	AC101367
C 39	74.2	24.7	76721	9	AL591024
C 40	74.2	24.7	209402	2	AC120348
C 41	74.2	24.7	264425	10	AC019028
C 42	73.6	24.5	5284	10	MUSRTKA
C 43	73	24.3	5892	10	ENUS3306
C 44	72	23.9	1040	5	AF007760
C 45	72	23.9	4360	10	AF402785

ALIGNMENTS

RESULT 1
AL138712/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-502P18 on chromosome 13,
complete sequence.
ACCESSION
AL138712
VERSION
AL138712.19
KEYWORDS
GI:15131448
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 164519)
AUTHORS
Bates,K.
TITLE
Direct Submission

JOURNAL

Submitted (03-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
UK. CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk

COMMENT

On Aug 9, 2001 this sequence version replaced g1:4715282. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as described above. All regions were finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as deletions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em1, EMBL; Sp1, SWISSPROT; Tr1, TREMBL; Wp1, WORMPEP. Information on the WORMPEP database can be found at

was obtained from <http://www.sanger.ac.uk/Projects/Celegans/wormmap> this sequence was generated from part of a cDNA clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13> R502p18 is from the library R5C11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/baccap/home.htm>

VECTOR: pBACe3 6

VECTOR: pBACE3.0
 IMPORTANT: This sequence is not the entire insert of clone Rpl11-502P18. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true right end of clone Rpl11-502P18 is at 164519 in this sequence. The true left end of clone Rpl11-956G is at 140266 in this sequence. The true right end of clone Rpl11-179F1 is at 23001 in this sequence.

REFERENCES

**WORKS
SOURCE**

Location/Qualifiers

```
Location/Qualifiers
1..164519
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-502P18"
/clone_id="8PCR11.2"
```

```

repeat_region      /CroneLib="RFL-III.2"  

repeat_region      216. .318  

repeat_region      /note="u6 repeat: matches 2. .104 of consensus"  

repeat_region      983. .1279  

repeat_region      /note="AlusX repeat: matches 1. .297 of consensus"  

repeat_region      1634. 1706  

repeat_region      /note="MER5A repeat: matches 95. .168 of consensus"  

repeat_region      1989. .2053  

repeat_region      /note="AluJ/FIAM repeat: matches 21. .86 of consensus"  

repeat_region      2724. .2892  

repeat_region      /note="L2 repeat: matches 2206. .2398 of consensus"  

repeat_region      3392. 3435  

repeat_region      /note="22 copies 2 mer ac 93% conserved"  

repeat_region      3453. 3743  

repeat_region      /note="AluJo repeat: matches 1. .291 of consensus"  

repeat_region      3982. 4146  

repeat_region      /note="MER5B repeat: matches 1. .161 of consensus"  

repeat_region      4495. 4586  

repeat_region      /note="FIAM_C repeat: matches 33. .124 of consensus"  

repeat_region      6847. 6934  

repeat_region      /note="44 copies 2 mer ca 75% conserved"  

repeat_region      8691. 9005  

repeat_region      /note="AluAb repeat: matches 1. .312 of consensus"  

repeat_region      10412. 10773  

repeat_region      /note="l1 repeat: matches 4244. .4621 of consensus"  

repeat_region      10814. 10964  

repeat_region      /note="l1 repeat: matches 3996. .4169 of consensus"  

repeat_region      11393. 11658  

repeat_region      /note="AlusG repeat: matches 1. .307 of consensus"  

repeat_region      12456. 12549

```

```

repeat_region
note="LIM4 repeat: matches 4612..4706 of consensus"
12550..12859
repeat_region
note="Alusg repeat: matches 1..310 of consensus"
12860..13131
repeat_region
note="LIM4 repeat: matches 4344..4612 of consensus"
13227..13760
repeat_region
note="LIM4 repeat: matches 5629..6165 of consensus"
13785..14015
repeat_region
note="LIM4 repeat: matches 5250..5477 of consensus"
14017..16233
repeat_region
note="LIM4 repeat: matches 4063..6299 of consensus"
16234..16506
repeat_region
note="Alub repeat: matches 30..301 of consensus"
16507..16680
repeat_region
note="LIM4 repeat: matches 3892..4063 of consensus"
16689..16849
repeat_region
note="LIM4 repeat: matches 5094..5255 of consensus"
16910..17043
repeat_region
note="LIM4 repeat: matches 4958..5103 of consensus"
17160..17462
repeat_region
note="Alusg repeat: matches 1..302 of consensus"
17794..17914
repeat_region
note="LIMEC repeat: matches 1622..1739 of consensus"
17915..18088
repeat_region
note="Alub repeat: matches 144..1295 of consensus"
18089..18379
repeat_region
note="Alusg repeat: matches 1..292 of consensus"
18380..18516
repeat_region
note="Alub repeat: matches 1..144 of consensus"
18517..18634
repeat_region
note="LIMEC repeat: matches 1495..1622 of consensus"
18637..18846
repeat_region
note="105 copies 2 mer tt 57% conserved"
18990..19900
repeat_region
note="Alusg repeat: matches 1..311 of consensus"
19902..19923
repeat_region
note="11 copies 2 mer tt 100% conserved"
21313..21486
repeat_region
note="FM repeat: matches 1..175 of consensus"
21878..22123
repeat_region
note="MIR repeat: matches 1..251 of consensus"
24957..25516
misc_feature
note="Sequence from clone PCR only."
25177..25343
repeat_region
note="Charlie3 repeat: matches 48..1215 of consensus"
25374..25537
repeat_region
note="82 copies 2 mer ta 61% conserved"
25538..25830
repeat_region
note="Alusg repeat: matches 1..266 of consensus"
26137..26754
repeat_region
note="MER2 repeat: matches 2..344 of consensus"
26845..27004
repeat_region
note="FM repeat: matches 1..160 of consensus"
27014..27230
repeat_region
note="Alusg/x repeat: matches 89..305 of consensus"
27233..27286
repeat_region
note="Alusg/x repeat: matches 79..133 of consensus"
27339..27650
repeat_region
note="Alusg repeat: matches 1..313 of consensus"
27679..27810
misc_feature
note="Single clone region. Sequence from reads from a short insert library derived from a single PUC clone. Restriction digest data confirm the assembly."
27828..28123
repeat_region
note="Alusg repeat: matches 1..259 of consensus"
29264..29564
repeat_region
note="Alub repeat: matches 1..301 of consensus"
29773..29973
repeat_region
note="MIR repeat: matches 37..235 of consensus"
31059..31337
repeat_region
note="Alub repeat: matches 7..289 of consensus"
31971..32253
repeat_region
note="Alusg repeat: matches 2..297 of consensus"

```

/note="AluX repeat: matches 1.295 of consensus"
 23264.29564
 /note="AluJ repeat: matches 1.361 of consensus"
 29773.29573
 /note="MIR repeat: matches 37.235 of consensus"
 31059.31337
 /note="AluB repeat: matches 7.289 of consensus"
 31971.31253
 /note="AluG repeat: matches 2.297 of consensus"

According to the Pre Publication Rules, every patent application received by the United States Patent and Trademark Office after November 29, 2000 will be pre-published at eighteen months from the effective filing date. When the application is published the contents, including the sequences, will become prior art.

Two new databases have been created to hold the pre-published sequences:

Published_Applications_NA contains nucleic acid sequences; the search results will have the extension **.rnpb**.

Published_Applications_AA contains amino acid sequences; the search results will have the extension **.rapb**.

Each pre-published application is given a unique Publication Number. An example of a Publication Number is US20021234567A1. The "US" indicates the application was a U.S. application. The first 4 digits show the calendar year the application was published. The next 7 digits represent when the application was published. This 7-digit number starts at zero at the beginning of each calendar year. Each application published is given the next number in order. The "A" indicates a utility patent application and the "1" shows that this was the first time the application had been published. If the applicants submit changes to the application, they may request that the changed application be published again. In such instances, the "1" at the end of the number would be replaced by a "2".

Sequences in the PGPub database are public information; it is permissible to leave these results in the case.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 11:59:28 ; Search time 52.5871 Seconds
(without alignments)
5977.502 Million cell updates/sec

Title: D64016_COPY_500_1440

Perfect score: 941

Sequence: 1 GCGGAAAGACAGCGACAC.....GCGGGGCTGTCGCGCGGG 941

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 350425 seqs, 19496369 residues 700850

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	52.6	5.6	1064	10	US-09-804-682-29
2	51.4	5.5	1919	10	US-09-814-777A-17
3	50.6	5.4	43058	10	US-09-954-456-292
4	50.6	5.4	43058	10	US-09-954-456-292
5	50.6	5.4	43058	10	US-09-954-456-292
6	48.4	5.1	42999	9	US-09-880-107-3950
7	48.4	5.1	42999	9	US-09-759-462A-17
8	47	5.0	2561	12	US-10-023-523-48
9	47	5.0	2561	12	US-10-023-523-48
10	46.4	4.9	2481	10	US-09-894-968-35
11	45	4.8	1614	12	US-10-023-523-45
12	45	4.8	1614	12	US-10-023-523-45
13	45	4.8	9968	10	US-09-764-877-2718
14	45	4.8	12425	12	US-10-023-523-50
15	45	4.8	12425	12	US-10-023-523-50
16	44.8	4.8	8459	10	US-09-817-913-8
17	44.8	4.8	8459	10	US-09-817-913-8
18	44.4	4.7	43058	10	US-09-954-456-292
19	44.4	4.7	43058	10	US-09-954-456-292

SUMMARIES

ALIGNMENTS

```

RESULT 1
US-09-804-682-29/c
Sequence 29, Application US/09804682
Patent No. US20020106763A1
GENERAL INFORMATION:
APPLICANT: Kinders, Robert
APPLICANT: Corey, Michael J.
TITLE OF INVENTION: PAL-18 POLYPEPTIDES, NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING THE SAME AND METHODS FOR SCREENING FOR OR
TITLE OF INVENTION: MODULATING THE SAME
FILE REFERENCE: 130001.406
CURRENT APPLICATION NUMBER: US/09/804.682
NUMBER OF SEQ ID NOS: 174
SCFIDARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 29
LENGTH: 1064
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 5, 6, 16, 21, 24, 25, 33, 39, 72, 110, 205, 214, 231, 232,
LOCATION: 235, 237, 238, 244, 245, 246, 256, 282, 252, 237, 306, 319,
LOCATION: 321, 323, 330, 334, 340, 349, 354, 355, 363, 372, 376, 378,
LOCATION: 397, 405, 432, 437, 454, 455, 457, 458, 459, 468, 470
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: 485, 487, 488, 494, 496, 499, 511, 524, 527, 552, 557, 562,
LOCATION: 585, 600, 611, 613, 623, 624, 652, 654, 674, 681, 687, 691,
LOCATION: 694, 701, 713, 716, 720, 721, 725, 731, 734, 735, 736, 743,
LOCATION: 744, 781, 782, 785, 789, 799, 803, 821, 823, 847, 852
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc feature
LOCATION: 858, 878, 884, 886, 896, 897, 901, 917, 920, 932, 939, 948,
LOCATION: 957, 961, 965, 981, 991, 993, 1001, 1002, 1005, 1011, 1018,
LOCATION: 1043, 1047, 1049, 1051, 1054, 1056
OTHER INFORMATION: n = A,T,C or G
US-09-804-682-29
Query Match 5.6%; Score 52.6; DB 10; Length 1064;
Best Local Similarity 41.5%; Pred. NC. 0.0047;
Matches 15; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

```

```

C 20 44.4 4.7 43058 10 US-09-880-107-3950
C 21 44.2 4.7 2236 9 US-09-978-295A-399
C 22 44.2 4.7 2236 9 US-09-978-697-399
C 23 44.2 4.7 2236 9 US-09-978-192A-399
C 24 44.2 4.7 2435 10 US-09-823-038A-40
C 25 43.4 4.6 8459 10 US-09-817-913-8
C 26 43.4 4.6 8459 10 US-09-817-913-8
C 27 43.4 4.6 152331 9 US-10-095-407-16
C 28 43.2 4.6 6314 10 US-09-954-456-1230
C 29 43.2 4.6 6314 10 US-09-954-456-1230
C 30 42.8 4.5 15500 10 US-09-764-869-2109
C 31 42.6 4.5 3110 10 US-09-764-877-3912
C 32 42.6 4.5 3110 10 US-09-764-877-3912
C 33 42.6 4.5 4689 9 US-09-860-846-34
C 34 42.6 4.5 4689 9 US-09-860-846-34
C 35 42.6 4.5 36778 9 US-09-860-846-34
C 36 42.6 4.5 36778 9 US-09-860-846-34
C 37 42.6 4.5 174424 10 US-09-967-768A-314
C 38 42.4 4.5 1065 10 US-09-804-682-33
C 39 42.2 4.5 1867 9 US-10-044-716-7
C 40 42.2 4.5 3547 9 US-10-044-716-7
C 41 42.2 4.5 4348 10 US-09-879-445-1
C 42 42.2 4.5 8147 12 US-10-109-886-9
C 43 42.2 4.5 152331 9 US-10-095-407-16
C 44 42 4.5 242 10 US-09-983-965-4700
C 45 42 4.5 5025 10 US-09-960-253-176

```

460 CGTCCGGGGCCGGGTGCGCAGCACCTCCCAAGCGCGGTGCGCCGGGCGCACCGGC 519

```

Query Match          5.4%; Score 50.6; DB 10; Length 43058;
Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 228; Conservative 0; Mismatches 259; Idents 2; Gaps 2;

QY 453 GGGCGCGCGTCGCCGGCCGCGTCGCACACACTCCGACGCGCTCGGCCCGGCC 512
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25300 GAGCGCCGCTGCTGGAGGCTCAGGGCCCCCGCCCGCTCCGAGGCTTCCTC-AGCC 25358
      | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Query Match 5.4%; Score 50.6; DB 10; Length 43058;
Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 228; Conservative 0; Mismatches 255; Indels 2; Gaps 2;
QY 513 ACCGCGCTCGTCCGCCCCCGCCCTCTCCGTAGCCAGGAGCGACCTCGGAGAA 572
DB 25359 ACCGCGAGCCAGCACTACTTCCCGC-ACITCGACCTGACCCGGGTCCGCGCACTIG 25417
QY 573 GAAGAGGTAGTGGGAGCGGATGAGGGGTGGGGACCCCTTGACGTACCAAGAGGA 632
DB 25418 CGCGCGACCGCTCCAAAGTGTGGCGCGGTGGGCGACCGGTGAAGAGCATCACAC 25477
QY 633 GTTGGCGGGTAGAAGTGGGTGGGAGAGGTATTAATCGCCCGCCCTCGGCTGCT 692
DB 25478 ATCGCGCGGCCCTGTCCAAAGTGTAGCGAGTGCACGCCCTACATCTCTGGCGTGCACCG 25537
QY 693 CTTCATCGAGTTCGCGGAGGCTCGAGCGCGCCGACACACTCTCTCTCGGCTCTC 752
DB 25538 GTCACTTCAAGGTGCGCGGGCGGTGCGGGCGGGCGGGCGGGCGGGCGGG 25597
QY 753 CCGGCGAGCGCGGCTCGAGCGGGCTCCGGGGTTCGAGCGCGGCGAGCGGCG 812
DB 25598 CGGGCGCGCGGGCGGGTTCGCGGGCGGGTGGGGTGGGGTGGGGTGGGGTGGGGT 25657
QY 813 CCGGCGCGGAGGATTACCGGGGAGGTGTGTCTCTCGCTGAGCGCGCGAGAGCGGC 872
DB 25658 CGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 25717
QY 873 GCTCAGCG 932
DB 25718 CGGGCGGGGAGGGGTGGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 25777
QY 933 GCGCGCGGG 941
DB 25778 GGTCCGCGG 25786

RESULT 4

US-09-954-456-529
; Sequence 529, Application US/0954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cano
; FILE REFERENCE: 589290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 529
; LENGTH: 43058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-529

Query Match 5.4%; Score 50.6; DB 10; Length 43058;
Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 228; Conservative 0; Mismatches 255; Indels 2; Gaps 2;
QY 453 GGGGCGCGCTCCGCGCGCGCGCTCCGAGCAGCTTCCCGAGCGCGCTCGGCGCGGCC 512
DB 25360 GAGGCGCGCGCTCCGAGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 25358
QY 513 ACCGCGCTCGTCCGCCCCCGCCCTCTCCGTAGCCAGGAGCGAGCGAGCGAGCGAG 632
DB 25359 ACCGCGAGCCAGCACTACTTCCCGC-ACITCGACCTGACCCGGGTCCGCGCACTIG 25417
QY 573 GAAGAGGTAGTGGGAGCGGATGAGGGGTGGGGACCCCTTGACGTACCAAGAGGA 572
DB 25418 CGCGCGACCGCTCCAAAGTGTGGCGCGGTGGGCGACCGGTGAAGAGCATCACAC 25477
QY 633 GTTGGCGGGTAGAAGTGGGTGGGAGAGGTATTAATCGCCCGCCCTCGGCTGCT 692
DB 25478 ATCGCGCGGCCCTGTCCAAAGTGTAGCGAGTGCACGCCCTACATCTCTGGCGTGCACCG 25537
QY 693 CTTCATCGAGTTCGCGGAGGCTCGAGCGCGCCGACACACTCTCTCTCGGCTCTC 752
DB 25538 GTCACTTCAAGGTGCGCGGGCGGTGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 25597
QY 753 CCGGCGAGCGCGGCTCGAGCGGGCTCCGGGGTTCGAGCGCGGCGAGCGGCG 812
DB 25598 CGGGCGCGCGGGCGGGTTCGCGGGCGGGTGGGGTGGGGTGGGGTGGGGTGGGGT 25657
QY 813 CCGGCGCGGAGGATTACCGGGGAGGTGTGTCTCTCGCTGAGCGCGCGAGAGCGGC 872
DB 25658 CGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 25717
QY 873 GCTCAGCG 932
DB 25718 CGGGCGGGGAGGGGTGGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 25777
QY 933 GCGCGCGGG 941
DB 25778 GGTCCGCGG 25786

RESULT 5

US-09-880-107-3950
; Sequence 3950, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3950
; LENGTH: 43058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 284721
US-09-880-107-3950

Query Match

Best Local Similarity 5.4%; Score 50.6; DB 10; Length 43058;
Matches 228; Conservative 0; Mismatches 255; Indels 2; Gaps 2;
QY 453 GGGGCGCGCTCCGCGCGCGCGCTCCGAGCAGCTTCCCGAGCGCGCTCGGCGCGGCC 512

RESULT 6
US-09-799-462A-17/c
Sequence 17, Application US/09799462A
Patent No. US20020160970A1
GENERAL INFORMATION:
APPLICANT: Haolaczky, Gyula
Szalay, Aladar
TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Heller Ehrman White & McCulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/799,462A
FILING DATE: 10-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,682
FILING DATE: 10-APR-1997
APPLICATION NUMBER: 08/695,191
FILING DATE: 07-AUG-1996
APPLICATION NUMBER: 08/682,060
FILING DATE: 15-JUL-1996
APPLICATION NUMBER: 08/829,822
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.

RESULT 7
 US-10-125-767-17/C
 : Sequence 17, Application US/10125767
 : Patent No. US20020160411A1
 : GENERAL INFORMATION:
 : APPLICANT: Hadjitzky, Gyula
 : TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND
 : METHODS
 : FOR PREPARING ARTIFICIAL CHROMOSOMES
 : NUMBER OF SEQUENCES: 34
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Heller Ehrman White & McCauliffe LLP
 : STREET: 4350 La Jolla Village Drive, 7th floor
 : CITY: San Diego
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 92122
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq Version 1.5
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/10/125,767
 : FILING DATE: 17-Apr-2002
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 09/724,693
 : FILING DATE: 28-NOV-2000
 : APPLICATION NUMBER: 08/835,682
 : FILING DATE: 10-APR-1997
 : APPLICATION NUMBER: 08/695,191
 : FILING DATE: 07-AUG-1996

Matches 144; Conservative 0; Mismatches 125; Indels 2; Gaps 2;

QY 673 CGCCCCCGCTCGGCTCTTCATCGAGTCCGGGGAGGCTCGAGCGCGGCGAGCG 732
DB 899 CGCGCGCGGAGCGGCGGCTCTCCCGCGCGGCGGCGGCGGCGGCGGCGGCG 840
QY 733 -GACACTCTCTCGGCTCTCTCCCGCGGCGGCTCGAGCGGCTCGGCGGCG 791
DB 839 CGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
QY 792 GGTCTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 851
DB 779 GGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 720
QY 852 GGTCTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 910
DB 719 GGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 910
QY 911 GACTCTGCGGCGGCGGCGGCTTTTGGCGGCGG 941
DB 659 GGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 629

RESULT 10
US-09-894-998-35/C
; Sequence 35, Application US/09894998
; Patent No. US20020090610A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: David C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998
; PRIORITY FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 2481
; TYPE: DNA
; ORGANISM: HSW-2
US-09-894-998-35

Query Match 4.9%; Score 46.4; DB 10; Length 2481;
Best Local Similarity 55.6%; Pred. No. 0.11;
Matches 89; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

DB 1341 CGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 764
QY 765 GCGCGGCTCGGAGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 824
DB 1281 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 824
QY 825 GATTACCGGGAAGTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1222
DB 1221 CCGCGAGGACTGTCAGACGCTCCCGCGGCGGCGGCGGCGGCGGCGGCG 1182

RESULT 11
US-10-023-529-45/C
; Sequence 45, Application US/10023529
; Patent No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING

; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: A. HEROSGLEROSIS
; CURRENT APPLICATION NUMBER: US/10/023,529
; PRIORITY FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1997-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1614)
US-10-023-529-45

Query Match 4.8%; Score 45; DB 12; Length 1614;
Best Local Similarity 46.9%; Pred. No. 0.21;
Matches 206; Conservative 0; Mismatches 230; Indels 3; Gaps 2;

QY 464 CCGCGGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 523
DB 730 CCGCGGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 671
QY 524 TCGGC--CCCGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 581
DB 670 GCTCTCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 611
QY 582 AGTGGGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 640
DB 610 CCGCGGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 551
QY 641 GGTAGGAAGTGGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 700
DB 550 GCGCGGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 491
QY 701 AGTTCGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 760
DB 490 GCGCGGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGG 431
QY 761 CCGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 820
DB 430 CCGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 371
QY 821 CGAGGATTACCGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 880
DB 370 CCGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 311
QY 881 CCGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 899
DB 310 GCGCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCGGCTCGGCG 292

RESULT 12
US-10-023-523-45/C
; Sequence 45, Application US/10023523
; Patent No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING

;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; FILE OF INVENTION: ATHEROSCLEROSIS
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/023,523
;; CURRENT FILING DATE: 2001-12-17
;; PRIOR APPLICATION NUMBER: US/09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US/09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US/08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US/60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US/60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 45
;; LENGTH: 1614
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(1614)
US-10-023-523-45

Query Match 4.8%; Score 45; DB 12; Length 1614;
Best Local Similarity 46.9%; Pred. No. 0.21;
Matches 206; Conservative 0; Mismatches 230; Indels 3; Gaps 2;
QY 464 CCGGGGCGCGCTCGCCACGACCTCCACAGCGCGCTCGGGCCGGGCGCCACCGCCCTCG 523
DB 730 CCGCGCGCCCGACCGCGCCCGCCCTCGCGCGCGCGCTGCGCTCGCGCGCGCGCGCTGCT 671
QY 524 TGGGC--CCCGCGCCCTCTCCCTAGCCGCGAGGAGCGAGCGCTGGAGGAGAGAGGCT 581
DB 670 CTCTGTGCGCGCGCGCGCGCTGTGCGCGCGCGCGCGAGCGCGCGCTCCCGCGCGCGCA 611
QY 582 AGCTGGGAGCGCGATGAGGCTGGGGAGCCCTTGACGTACACAGAGAG--GTGCCGG 640
DB 610 CCG 551
QY 641 GGTAGGAGTGGCGTGGGAGAGGTTTAAATCGCCCGCGCGCTCGCTGCTCTCTCTCTCT 700
DB 550 CG 451
QY 701 AGCTTCGCGGAGGCTTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 760
DB 490 GCGCAGCGAGGCGCGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 431
QY 761 CCGCGCGCGCTCGGAGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 820
DB 430 CG 371
QY 821 CGAGGATTACCGCGGGAAGTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 880
DB 370 CG 311
QY 881 CG 899
DB 310 CCGCGCGCGCGCTGGACGCG 292

RESULT 13
US-09-764-877-2718
;; Sequence 2718, Application US/09764877
;; Patent No. US20020147140A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC005
;; CURRENT APPLICATION NUMBER: US/09/764,877
;; CURRENT FILING DATE: 2001-01-17

;; Prior application data removed - refer to PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 4031
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2718
;; LENGTH: 9968
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-764-877-2718
Query Match 4.8%; Score 45; DB 10; Length 9968;
Best Local Similarity 45.8%; Pred. No. 0.27;
Matches 232; Conservative 0; Mismatches 270; Indels 5; Gaps 2;
QY 387 CTGAGCGGATGGGCGAGGAGGCGGCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 446
DB 19 CTCGCGCGCTGCTGCG 78
QY 447 CCG 506
DB 79 TCAGGCGGCG 138
QY 507 CGGCG 566
DB 139 GCGCTGCG 198
QY 567 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 626
DB 199 CCG 258
QY 627 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 586
DB 259 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 317
QY 687 CT 746
DB 318 GCG 377
QY 747 CT 806
DB 378 CCG 433
QY 807 CCGCGCGCGCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 866
DB 434 TGGTGCG 493
QY 867 ACGGGCGCTCAGGCG 893
DB 494 CCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 520

RESULT 14
US-10-023-529-50/c
;; Sequence 50, Application US/10023529
;; Patent No. US20020129386A1
;; GENERAL INFORMATION:
;; APPLICANT: Lees, Ann M.
;; APPLICANT: Lees, Robert S.
;; APPLICANT: Law, Simon W.
;; APPLICANT: Arjona, Anibal A.
;; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
;; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
;; TITLE OF INVENTION: ATHEROSCLEROSIS
;; FILE REFERENCE: 10797-004001
;; CURRENT APPLICATION NUMBER: US/10/023,529
;; CURRENT FILING DATE: 2001-12-17
;; PRIOR APPLICATION NUMBER: 09/616,289
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/517,849
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: December 8, 2002, 09:24:02 : Search time 37,4297 Seconds
(without alignments)
7710.006 Million cell updates/sec

Title: D64016_COPY_500_1440
Per set score: 941
Sequence: 1 GCCGAAAGACAGGACAC.....GCCGGGTCTTGGCGGG 941

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Sequences: .441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Data base: Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Res	ID	Score	Query Match	Length	DB ID	Description
1	941	100.0	1745	2	US-08-556-424-1	Sequence 1, Appl
2	941	100.0	1745	2	US-09-263-914-1	Sequence 1, Appl
3	194.8	20.7	2313	1	US-08-232-538-5	Sequence 5, Appl
4	194.8	20.7	2313	1	US-09-427-553-1	Sequence 1, Appl
5	194.8	20.7	2523	4	US-09-051-363-1	Sequence 1, Appl
6	194.8	20.7	2651	2	US-08-786-164-5	Sequence 5, Appl
7	55.4	5.9	15378	3	US-08-785-420-1	Sequence 1, Appl
8	53.2	5.7	1785	3	US-08-729-416C-16	Sequence 16, Appl
9	51.4	5.5	7218	1	US-08-232-463-14	Sequence 1, Appl
10	50.6	5.4	801	4	US-09-298-568-1	Sequence 3, Appl
11	48.6	5.2	4523	4	US-09-473-716-1	Sequence 1, Appl
12	48.2	5.1	3602	2	US-08-883-795A-33	Sequence 1, Appl
13	48.2	5.1	3602	4	US-09-018-138-1	Sequence 3, Appl
14	47.4	5.0	1335	5	PCT-US91-06532-1	Sequence 1, Appl
15	47.2	5.0	5703	4	US-08-280-590A-36	Sequence 36, Appl
16	46.8	5.0	1327	4	US-08-483-533-36	Sequence 36, Appl
17	46.8	5.0	1327	4	US-09-283-471A-36	Sequence 36, Appl
18	46.6	5.0	7898	4	US-08-984-709A-49	Sequence 4, Appl
19	46	4.9	46	2	US-08-556-424-4	Sequence 4, Appl
20	46	4.9	46	4	US-09-263-914-4	Sequence 4, Appl
21	46	4.9	46	4	US-09-017-631-23	Sequence 23, Appl
22	45.8	4.9	6453	3	US-08-306-591B-14	Sequence 14, Appl
23	45.8	4.9	6453	3	US-09-209-568-10	Sequence 10, Appl
24	45.8	4.9	6453	3	US-09-356-952-8	Sequence 8, Appl
25	45.8	4.9	9515	1	US-08-920-812-13	Sequence 13, Appl
26	45.8	4.9	9515	1	US-08-920-827-13	Sequence 13, Appl
27	45.8	4.9	9515	1	US-08-921-177-13	Sequence 13, Appl

C 28	45.8	4.9	9515	1	US-08-362-577C-13	Sequence 13, Appl
C 29	45.8	4.9	9515	2	US-08-920-828-13	Sequence 13, Appl
C 30	45.6	4.8	44377	2	US-08-804-227C-7	Sequence 7, Appl
C 31	45.6	4.8	44377	2	US-08-804-198-1	Sequence 1, Appl
C 32	45.4	4.8	4524	2	US-08-845-988-7	Sequence 7, Appl
C 33	45.4	4.8	4524	3	US-09-206-537-7	Sequence 7, Appl
C 34	45.4	4.8	4524	4	US-09-430-834-7	Sequence 7, Appl
C 35	45.4	4.8	4791	4	US-08-949-155-49	Sequence 49, Appl
C 36	45.4	4.8	4791	4	US-09-819-964-49	Sequence 49, Appl
C 37	45.4	4.8	43280	2	US-08-804-227C-1	Sequence 1, Appl
C 38	45.4	4.8	169998	4	US-09-676-610B-24	Sequence 24, Appl
C 39	45.2	4.8	224	3	US-09-017-631-24	Sequence 24, Appl
C 40	45.2	4.8	224	4	US-09-018-138-2	Sequence 2, Appl
C 41	44.8	4.8	1896	4	US-09-343-011B-4	Sequence 4, Appl
C 42	44.8	4.8	1896	4	US-09-343-011B-4	Sequence 4, Appl
C 43	44	4.7	801	2	US-08-770-379-16	Sequence 16, Appl
C 44	44	4.7	801	4	US-08-757-669A-16	Sequence 16, Appl
C 45	44	4.7	801	4	US-09-230-371A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-556-424-1
: Sequence 1, Application US/08556424
: Patent No. 5916753
: GENERAL INFORMATION:
: APPLICANT: Williams, Lewis T.
: APPLICANT: Morishita, Kaoru
: TITLE OF INVENTION: No. 5916753el Promoter for VEGF Receptor
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSES: Townsend and Townsend and Crew
: STREET: One Market Plaza, Steuart Tower, Suite 2000
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/556,424
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Dow, Karen B.
: REGISTRATION NUMBER: 29,684
: REFERENCE/DOCKET NUMBER: 2307K-0635000US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-543-5600
: TELEFAX: 415-543-5043
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1745 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: Single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: 1..1155
: OTHER INFORMATION: /note= "Nucleotides numbered 1
: OTHER INFORMATION: through 1195 correspond to -1195 through -1 from
: OTHER INFORMATION: Figure 2."
US-08-556-424-1

Query Match 100.0% Score 941; DB 2; Length 1745;
Best Local Similarity 100.0%; Pred. No. 1.6e-191;
Matches 941; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 60
DB 500 GCGGAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 559
QY 61 GTCCCAAGCAGCGCTCAGTTCCTCCCTCAGCGCTCCAGCTTCAGTTCCTTGTCCGAGGGT 120
DB 560 GTCCCAAGCAGCGCTCAGTTCCTCCCTCAGCGCTCCAGCTTCAGTTCCTTGTCCGAGGGT 619
QY 121 CTCGGTGCCTTCAGCTTCCTGGGACAGTCTGAAGGGGTGAGGGGTGAGGGGTGAGGGGT 180
DB 620 CTCGGTGCCTTCAGCTTCCTGGGACAGTCTGAAGGGGTGAGGGGTGAGGGGTGAGGGGT 679
QY 181 GCGGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 240
DB 680 GCGGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 739
QY 241 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 740 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 799
QY 301 GCGGAGCTTCAGCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 360
DB 800 GCGGAGCTTCAGCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 559
QY 361 GGTGAGGAGCTTCGACGATTCCTGAGCGGCTGGGACGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 860 GGTGAGGAGCTTCGACGATTCCTGAGCGGCTGGGACGAGGAGGAGGAGGAGGAGGAGGAGG 919
QY 421 AGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 920 AGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 579
QY 481 AGCACCCTCCACGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 540
DB 980 AGCACCCTCCACGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1039
QY 541 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 1040 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1099
QY 601 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 1100 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1159
QY 661 AAGGTTAATAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 1160 AAGGTTAATAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1219
QY 721 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
DB 1220 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1279
QY 781 CTCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 840
DB 1280 CTCGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 1339
QY 841 GGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 900
DB 1340 GGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 1399
QY 501 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 941
DB 1400 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

RESULT 2

US-09-263-914-1

; Sequence 1, Application US/09263914

; Patent No. 6245512

; GENERAL INFORMATION:

; APPLICANT: Williams, Lewis T.

; APPLICANT: Morishita, Kaoru

;; TITLE OF INVENTION: No. 6245512el Promoter for VEGF Receptor
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew
;; STREET: One Market Plaza, Stewart Tower, Suite 2000
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/263-914
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/556-424
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dow, Karen S.
;; REGISTRATION NUMBER: 29,684
;; REFERENCE/DOCKET NUMBER: 2307K-0635000S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-543-9600
;; TELEFAX: 415-543-5043
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1745 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: 1..1195
;; OTHER INFORMATION: /note= "Nucleotides numbered 1
;; OTHER INFORMATION: through 1195 correspond to -1195 through -1 from
;; US-09-263-914-1

Query Match 100.0%; Score 941; DB 4; Length 1745;
Best Local Similarity 100.0%; Pred. No. 1.6e-191;

Matches 941; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 60
DB 500 GCGGAAAGACGACGACGCTCCCTGGGACCTGAGCTGGTTCAGCTCTTCCCAAG 559
QY 61 GTCCCAAGCAGCGCTCAGTTCCTCCCTCAGCGCTCCAGCTTCAGTTCCTTGTCCGAGGGT 120
DB 560 GTCCCAAGCAGCGCTCAGTTCCTCCCTCAGCGCTCCAGCTTCAGTTCCTTGTCCGAGGGT 619
QY 121 CTCGGTGCCTTCAGCTTCCTGGGACAGTCTGAAGGGGTGAGGGGTGAGGGGTGAGGGGT 180
DB 620 CTCGGTGCCTTCAGCTTCCTGGGACAGTCTGAAGGGGTGAGGGGTGAGGGGTGAGGGGT 679
QY 181 GCGGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 240
DB 680 GCGGGAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 739
QY 241 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 740 ACCGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 799
QY 301 GCGGAGCTTCAGCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 360
DB 800 GCGGAGCTTCAGCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 859
QY 361 GGTGAGGAGCTTCGACGATTCCTGAGCGGCTGGGACGAGGAGGAGGAGGAGGAGGAGGAGG 420

Db 860 GGGTGGAGGAGTCTGCAAGGATTTCCTAGCCGATGGGAGGAGGAGGCGCA 919
QY 421 AGAGGCGGAGCAACACCTCAACCTCCCGGCGCGCTCCCGGCGCGGCGGCC 480
Db 920 AGAGGCGGAGCAACACCTCAACCTCCCGGCGCGCTCCCGGCGCGGCGGCC 979
QY 481 AGCACTTCCCGGCGGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db 980 AGCACTTCCCGGCGGCTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1039
QY 541 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 1040 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
QY 601 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 1100 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1159
QY 661 AAGTTATATAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 1160 AAGTTATATAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1219
QY 721 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 1220 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1279
QY 781 CTCGGGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Db 1280 CTCGGGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1339
QY 841 GGTGTCTCTCTGCTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 1340 GGTGTCTCTCTGCTGGAGCGGAGAGCGGCGGCTCAGGCGGCGGCGGCGGCGG 1399
QY 901 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 941
Db 1400 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

RESULT 3

US-08-232-538-5
; Sequence 5, Application US/08232538
; Patent No. 5712380
; GENERAL INFORMATION:
; APPLICANT: THOMAS, KENNETH A.
; APPLICANT: KENDALL, RICHARD L.
; TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL CELL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000 126 E Lincoln Avenue
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,538
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen, John W.III
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 188881A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 5:
; LENGTH: 2313 base pairs
; TYPE: nucleic acid

SEQUENCE CHARACTERISTICS:
LENGTH: 2313 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-232-538-5

Query Match 20.7%; Score 194.8; DB 1; Length 2313;
Best Local Similarity 98.1%; Pred. No. 3.7e-33;
Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 730 GCGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCGGCGGCGGCGGCGG 789
Db 1 GCGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCGGCGGCGGCGGCGGCG 60
QY 790 TCGGGTGCAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 849
Db 61 TCGGGTGCAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 118
QY 850 CTGGCTGGAGCGCGGAGCGGCGGCGGCTCAGGGCGGCGGCGGCGGCGGCGG 909
Db 119 CTGGCTGGAGCGCGGAGCGGCGGCGGCTCAGGGCGGCGGCGGCGGCGGCGG 178
QY 910 CGGACTCTGCGGCGGCGGCTCTTGGCGGCGGCGG 941
Db 179 CGGACTCTGCGGCGGCGGCTCTTGGCGGCGGCGG 210

RESULT 4

US-09-427-353-1
; Sequence 1, Application US/09427353
; Patent No. 6375929
; GENERAL INFORMATION:
; APPLICANT: THOMAS, KENNETH A.
; APPLICANT: GOLDMAN, COREY K.
; APPLICANT: KENDALL, RICHARD L.
; APPLICANT: BETT, ANDREW J.
; APPLICANT: HUCKLE, WILLIAM R.
; TITLE OF INVENTION: GENE THERAPY FOR INHIBITION OF
; TITLE OF INVENTION: ANGIOGENESIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,353
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19810YCA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2313 base pairs
; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-09-427-353-1

```

Query Match 20.7%; Score 194.8; DB 4; Length 2313;
Best Local Similarity 98.1%; Pred. NO. 3.7e-33;
Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

730 GCGGACACTCCTCTCGGTCCTCCCCGGAGCGGGCGGGTCGGAGCGGGCTCCGGGGC 789
 QY
 1 GCGGACACTCCTCTCGGTCCTCCCCGGAGCGGGCGGGTCGGAGCGGGCTCCGGGGC 60
 DB

790 TCGGGTGCAGCGGCCAGCGGGCGCTGGCGCGAGGATTACCCGGGGAAGTGGTTGCTC 849

850 CTGGCTGGAGCGCGGAGACGGCGGCTCAGGGCGGGCGGGCGGCGGACGAGGGA 909

QY 910 CGGACTCTGGCGGCCGGGGTCTTTGGCCCGCGGG 941
|||||
Db 179 CGGACTCTGGCGGCCCGGGTCTTTGGCCCGCGGG 210

RESULT 5
 US-09-051-363-1
 Sequence 1, Application US/09051363
 Patent No. 6270993
 GENERAL INFORMATION:
 APPLICANT: Shibuya, Masabumi
 APPLICANT: Okamoto, Masaji
 APPLICANT: Niwa, Mikio
 APPLICANT: Matsumoto, Tomoe
 APPLICANT: Asano, Makoto
 APPLICANT: Segawa, Toshiaki
 TITLE OF INVENTION: VEGF-BINDING POLYPEPTIDES
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street

```

1  INFORMATION FOR SEQ ID NO: 1:
2  =====
3  SEQUENCE CHARACTERISTICS:
4  =====
5  LENGTH: 2523 base pairs
6  TYPE: nucleic acid
7  STRANDEDNESS: double
8  TOPOLOGY: linear
9  MOLECULE TYPE: cDNA
10  ORIGINAL SOURCE:
11  =====
12  ORGANISM: Homo sapiens
13  CELL TYPE: placental tissue
14  FEATURE:
15  =====
16  NAME/KEY: Coding Sequence
17  LOCATION: 250..2523
18  US-09-051-363-1

```

Query Match 20.7%; Score 194.8; DB 4; Length 2523;
Best local Similarity 98.1%; Pred. No. 3.7e-33;
Matches 208; Conservative 0; Mismatches 2; Indels 2; Gaps 1.

Qy 730 GCGGACACTCCTCTCGGTCCTCCCGGGCAGCGGCGGTCGGAGCGGCTCCGGGGC 789

QY 750 TCGGGTGCAGCGGCCACGCGGGCGCGTGGCGGCGAGGATTACCGCGGGAGTAGTGTTGTC 849

Qy 850 CTGGCTGAGCCCGAGACGGGGCGCTCAGGGCGGGGGCGGGCGGAGCAGAGCA 909

Qy	910	CGGAC	CTGGG	CGCGG	CTTTGG	CGCGGG	941
Db	179	CGGAC	CTGGG	CGCGG	CTGTTGG	CGCGGG	210

RESULT 5
US-08-786-164-5
: Sequence 5, Application US/08786164
: Patent No. 5861484
: GENERAL INFORMATION:
: APPLICANT: THOMAS, KENNETH A.
: APPLICANT: KENDALL, RICHARD L.
: TITLE OF INVENTION: INHIBITOR OF VASCULAR ENDOTHELIAL
: TITLE OF INVENTION: CELL GROWTH FACTOR
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: P.O. Box 2000, 126 E. Lincoln Ave.
: CITY: Rahway
: STATE: NJ
: COUNTRY: USA
: ZIP: 07065-0900
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: Microsoft Word 6.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/786.164
: FILING DATE: 21-JAN-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Mark Hand, J
: REGISTRATION NUMBER: 36,545
: REFERENCE/DOCKET NUMBER: 18688DA
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 908-594-3905
: TELEFAX: 908-594-4720
: TELEFAX:

Dd	668	CGGGTCCCGCGGCGGGGGCGGGACCAAGCGCGGCCGCCTGTAGCGG	629
Qy	276	ACTTTCCGCCCTAAGSCCAGCGGGGGAGCTTCAGCCATTGCTCCATCCCCAGTTTCGGG	337
Dd	628	GCGCGCGCCCATGCCCGCGGGGAGCGCGCAGACC CGCGGCTCCCGCGCAGGGG	569
Qy	338	CGCCCCCAGACGTAGTAAGCCGGTGCGAGGGAATCTCGAACGATTCCTGTAGCGCGAT	397
Dd	568	GGCGGGGCGCCCGGGGATCGGGCGCGGCGCCCGGGGGTCCGGCGCCGCGCGAGA	509

Db 644 TACTACCGTGCACCCCATCTCCGGCTGCGCGGGTCCCTGACGACCTGGGGCGG 585
QY 726 CCAGCGGACACATCTCTCGGCTCTCCCGGAGCGCGGCTCGGAGCGGCTCCG 785
Db 584 CCGGTGACACACCGCGGCGCTCTCTCCCGGAGCGCGGCTCGGAGCGGCTCCG 525
QY 786 GGGCTCGGGTGCAGCGCGCGCGCTGCGCGGAGGATACCGCGGAGGTGGT 845
Db 524 CCGGCTAGAGGAGCGGCGCTGTCAGGGCGCGGCTGTCGAGGAGCGGCGCGG 455
QY 846 TCTCTGCTGGAGCGCGGAGCGGCGCTGTCAGGGCGCGGCTGTCGAGGAGCGGCGG 905
Db 464 TGGCGCGGCTCGGCTCGGCGGAGCGCGGCTGTCGAGGAGCGGCGGCGG 405
QY 906 AGGAC 910
Db 404 GTGGC 400

RESULT 13
09-018-138-1/c
Sequence 1, Application US/09018138
Patent No. 6187564
GENERAL INFORMATION:
APPLICANT: Sytkowski, Arthur J.
TITLE OF INVENTION: Production and Use of Recombinant
FILE REFERENCE: BH97-05A
CURRENT FILING DATE: US/09/018,138
EARLIER FILING DATE: 1998-02-03
EARLIER FILING DATE: 08/890,929
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3602
TYPE: DNA
ORGANISM: Human
US-09-018-138-1

Query Match 5.1%; Score 48.2; DB 4; Length 3602;
Best Local Similarity 49.8%; Pred. No. 0.05;
Matches 122; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
QY 666 TATAATCGCCCGCCCTCGGCTCTCTTCATCGAGGTCGCGGAGGCTCGGAGCGG 725
Db 644 TACTACCGTGCACCCCATCTCCGGCTGCGCGGGTCCCTGACGACCTGGGGCGG 585
QY 726 CCAGCGGACACATCTCTCGGCTCTCCCGGAGCGCGGCTCGGAGCGGCTCCG 785
Db 584 CCGGTGACACACCGCGGCGCTCTCTCCCGGAGCGCGGCTCGGAGCGGCTCCG 525
QY 786 GGGCTCGGGTGCAGCGCGCGCGCTGCGCGGAGGATACCGCGGAGGTGGT 845
Db 524 CCGGCTAGAGGAGCGGCGCTGTCAGGGCGCGGCTGTCGAGGAGCGGCGCGG 455
QY 846 TCTCTGCTGGAGCGCGGAGCGGCGCTGTCAGGGCGCGGCTGTCGAGGAGCGGCGG 905
Db 464 TGGCGCGGCTCGGCTCGGCGGAGCGCGGCTGTCGAGGAGCGGCGGCGG 405
QY 906 AGGAC 910
Db 404 GTGGC 400

RESULT 14
PCT-US91-06532-1/c
Sequence 1, Application PCT/US9106532
GENERAL INFORMATION:
APPLICANT: Roizman, Bernard
TITLE OF INVENTION: Recombinant Herpes Simplex Viruses
NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza Suite 2100
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent's Release #1.0, Version #1.25
CURRENT APPLICATION LAIR:
APPLICATION NUMBER: PCT/US91/06532
FILING DATE: 19910910
CLASSIFICATION: 42
ATTORNEY/AGENT INFORMATION:
NAME: Gruber, Lewis S.
REGISTRATION NUMBER: 30,060
REFERENCE/DOCKET NUMBER: 27373/5235
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1335 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US91-06532-1

Query Match 5.0%; Score 47.4; DB 5; Length 1335;
Best Local Similarity 46.0%; Pred. No. 0.067;
Matches 271; Conservative 0; Mismatches 311; Indels 7; Gaps 3;
QY 357 ACCCGGCTGAGGAGTCTCGAGGATTCCTAGGCGGATCGGAGGAGGAGGCGGCAAG 416
Db 946 ACCGGTGTGGTGGGAGCTCCCGGGAGCTCCGGGAGAGACCCAGGCGGCTCGGGTGTA 927
QY 477 GCGAAGAGGCGGCGGAGCAAGACCTCGGCGGCGGCGCTCCCGGCGGCGCGCG 476
Db 926 GGTATAGACGAGTTCGCGGCGGCGGCTCCCGGCGGCGGCGGCGGCGGCGGCGG 867
QY 477 GCGAGGACCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 536
Db 866 CCGAGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 808
QY 537 CTCCTCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 596
Db 807 GTCGCGCGGCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 748
QY 557 TACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 556
Db 747 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 688
QY 657 GGGAAAGGATTAATATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 716
Db 687 CCGGCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 628
QY 717 CGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 776
Db 627 GGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 568
QY 777 CGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 836
Db 567 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 509
QY 837 AAGTGTGTGTCTCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 896
Db 508 AGGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 449

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:50:37 ; Search time 189.314 Seconds
(without alignments)
11193.756 Million cell updates/sec

Title: D64016_COPY_500_1440
Perfect score: 941
Sequence: 1 GCCGAAAGACACGACAC.....GCCGGTCITTCGCCCGGG 941

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
tail number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
N_Geneseq_101002:*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*		
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	941	100.0	1745	18	Flt-1 VEGF receptor
2	667	70.9	840	17	Promoter of human
3	613	65.1	6316	22	Tumour suppressor
4	613	65.1	6316	24	Chemically treated
5	613	65.1	6316	24	Human gene regulat
6	613	65.1	6316	24	Signal transductio
7	613	65.1	6316	24	Human angiogenesis
8	510.8	54.3	6316	22	Tumour suppressor
9	510.8	54.3	6316	24	Chemically treated

C 10	510.8	54.3	6316	24	AA561171	Human gene regulat
C 11	510.8	54.3	6316	24	ABK31261	Signal transductio
C 12	510.8	54.3	6316	24	ABQ67064	Human angiogenesis
C 13	194.8	20.7	2313	19	AAV19347	Human soluble VEGF
C 14	194.8	20.7	2313	19	AAV09330	Soluble VEGF recep
C 15	194.8	20.7	2523	18	AAV01457	Human VEGF recepto
C 16	194.8	20.7	2523	18	AAV01457	Human VEGF recepto
C 17	194.8	20.7	2523	18	AAV01457	Human VEGF recepto
C 18	194.8	20.7	2523	18	AAV01457	Human VEGF recepto
C 19	194.8	20.7	2523	18	AAV01457	Human VEGF recepto
C 20	130.6	13.9	3009	22	AA591131	CDNA encoding amin
C 21	69	7.3	114955	20	AA591131	SVGF-R1 gene. Ho
C 22	64.2	6.8	114955	20	AA591131	Human soluble vasc
C 23	63.6	6.8	114955	20	AA591131	DNA encoding novel
C 24	59	6.3	1127	21	AA02477	Flt-1extraFASm/Cy
C 25	57.8	6.1	1059	21	AA01833	Human adenosine A1
C 26	56.2	6.0	154746	24	AA025519	Human gene express
C 27	56.2	6.0	154746	24	AA025519	Human colon cancer
C 28	55.6	5.9	600	24	AB052496	Human colon cancer
C 29	55.6	5.9	600	24	AB052496	Human herpesvirus
C 30	55.4	5.9	15377	13	AA025975	Oligonucleotide fo
C 31	55	5.8	1218	21	AA02488	MM mutant porcine
C 32	54.8	5.8	1000	21	AA02484	Human colon cancer
C 33	53.6	5.7	989	21	AA02539	Human colon cancer
C 34	53.2	5.7	1785	18	AA166990	Rat brain-specific
C 35	51.6	5.5	754	24	AB040800	Oligonucleotide fo
C 36	51.6	5.5	754	24	AB040800	Oligonucleotide fo
C 37	51	5.4	629	24	AB043450	Oligonucleotide fo
C 38	51	5.4	629	24	AB043450	Oligonucleotide fo
C 39	50.8	5.4	35465	22	AA54723	Nucleotide sequenc
C 40	50.6	5.4	801	21	AAA30291	Rhadinovirus cis-
C 41	50.6	5.4	801	21	AAA30291	Kaposi's sarcoma-a
C 42	50.6	5.4	43058	24	ABN97455	Gene #3953 used to
C 43	50.6	5.4	43058	24	ABN97455	Lung cancer relate
C 44	50.6	5.4	43058	24	ABN97455	Lung cancer relate
C 45	50.4	5.4	15672	12	AAQ10613	Rianodin receptor

ALIGNMENTS

RESULT 1
AA169249 standard, DNA, 1745 BP.

AA169249:
11-JAN-1998 (first entry)

Flt-1 VEGF receptor gene promoter region.

Flt-1: VEGF receptor; vascular endothelial growth factor; human;
vascular permeability factor; fms-like receptor tyrosine kinase;
hyperplasia; thrombosis; atherosclerosis; restenosis;
inflammatory disease; haemophilia; wound healing;
diabetic retinopathy; rheumatoid arthritis; blood cell trafficking;
tumour; angiogenesis; gene therapy; transgenic animal; ss.

Homo sapiens.

Key	Location/Qualifiers
promoter	447..1479
protein_bind	/tag= a /note= "(Claim 4)" 1051..1055
protein_bind	/tag= b /note= "E26 transformation specific sequence" 1068..1071
misc_feature	/tag= c /note= "E26 transformation specific sequence" 1098..1103


```
QY 776 GCGGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 835
Db 61 GCGGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 120
QY 836 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGCGGCGCGGCGG 895
Db 121 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGCGGCGCGGCGG 180
QY 896 GCGAAGCAGAGGAGCGGACTCTGCGGCGCGGCGGCTTGGCGGCGG 941
Db 181 GCGAAGCAGAGGAGCGGACTCTGCGGCGCGGCGGCTTGGCGGCGG 226

RESULT 6
AL543052 972 bp mRNA linear EST 16-FEB-2001
LOCUS AL543052 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1002Y109 5
DEFINITION prime, mRNA sequence.
ACCESSION AL543052
VERSION AL543052.1 GI:12875530
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 972)
Li.W.B., Gruber,C., Jessee,J., and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 972
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1002Y109"
/tissue_type="placenta"
/Note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life technologies. Contact : Feng Liang Life technologies,
a division of invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 259 a 248 c 266 g 194 t 5 others
ORIGIN
Query Match 23.7% Score 223.4; DB 9; Length 972;
Best Local Similarity 59.1%; Pred. No. 3.1e-33;
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 716 TCGGAGCGGGCGGCGGAGCGGACTCTCTCGGCTCTCCCGCGGCGGCGGCGGCTCGGA 775
Db 1 TCGGAGCGGGCGGCGGAGCGGACTCTCTCGGCTCTCCCGCGGCGGCGGCGGCTCGGA 50
QY 776 GCGGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 835
Db 61 GCGGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 120
QY 836 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGGCGGCGGCGGCG 895
Db 121 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGGCGGCGGCGGCG 180
QY 896 GCGAAGCAGAGGAGCGGACTCTGCGGCGCGGCGGCTTGGCGGCGG 941
Db 181 GCGAAGCAGAGGAGCGGACTCTGCGGCGCGGCGGCTTGGCGGCGG 226

RESULT 7
AL547815 1002 bp mRNA linear EST 16-FEB-2001
LOCUS AL547815 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1033YD17 5
DEFINITION prime, mRNA sequence.
ACCESSION AL547815
VERSION AL547815.1 GI:12882232
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1002)
Li.W.B., Gruber,C., Jessee,J., and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 1002
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1033YD17"
/tissue_type="placenta"
/Note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life technologies. Contact : Feng Liang Life technologies,
a division of invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 251 a 240 c 281 g 229 t 1 others
ORIGIN
Query Match 23.6% Score 221.8; DB 9; Length 1002;
Best Local Similarity 59.1%; Pred. No. 6.3e-33;
Matches 223; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 717 CGGAGCGCGGCGGAGCGGACTCTCTCGGCTCTCCCGCGGCGGCGGCGGCTCGGAG 776
Db 1 CGGAGCGCGGCGGAGCGGACTCTCTCGGCTCTCCCGCGGCGGCGGCGGCTCGGAG 60
QY 777 CGGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 836
Db 61 CGGGCTCCGGGCTCGGGTGCAGCGCCAGCGGCGCTTGGCGGCGAGGATTACCGGG 120
QY 837 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGGCGGCGGCGGCG 896
Db 121 GAAGTGGTGTCTCTGCTGAGCGCGGAGCGGCGCTCAGGCGCGGCGGCGGCGGCG 180
QY 897 GCGAAGCAGAGGAGCGGACTCTGCGGCGCGGCGGCTTGGCGGCGG 941
Db 181 GCGAAGCAGAGGAGCGGACTCTGCGGCGCGGCGGCTTGGCGGCGG 225

RESULT 8
AL136704 662 bp mRNA linear EST 02-AUG-2002
LOCUS AL136704 PLACE1 Homo sapiens cDNA clone PLACE1004901 5', mRNA
DEFINITION sequence.
ACCESSION AL136704
VERSION AL136704.1 GI:10997243
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 562)
Ota, T., Nishikawa, I., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, I., Sugano, S. and
Isogai, T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5' & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute
Location/Qualifiers
1. 662

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PLACE104901"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: pME18SF3"

BASE COUNT 153 a 175 c 199 g 121 t 14 others

Query Match 22.9%; Score 215.4; DB 9; Length 662;

Best Local Similarity 97.7%; Pred. No. 1.1e-31;

Matches 216; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 721 GCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGAGCGGG 780

Db 1 GCGGCCAGCGGACACTCTCTCGGCTCTCTCCCGGACGCGCGGCTCGAGCGGG 60

QY 781 CTCGGGGCTCGGTCAGCGGCGGCGGCGGCTCGGCGGCGGAGTACC 840

Db 61 CTCGGGGCTCGGTCAGCGGCGGCGGCGGCTCGGCGGCGGAGTACC 120

QY 841 GGTCTCTCTCGGTCAGCGGCGGCGGCGGCTCGGCGGCGGCGGCGGCGG 900

Db 121 GGTCTCTCTCGGTCAGCGGCGGCGGCGGCTCGGCGGCGGCGGCGGCGG 180

QY 901 ACGAGAGCGGACACTCTCGGCGGCGGCGGCTCTTGCCCGCGGG 941

Db 181 ACGAGAGCGGACACTCTCGGCGGCGGCGGCTCTTGCCCGCGGG 221

RESULT 9
AL551816 1025 bp mRNA linear EST 16-FEB-2001
DEFINITION
AL551816 L11_NFL006.PL2 Homo sapiens cDNA clone CS0D1061918.5
prime, mRNA sequence.

ACCESSION AL551816

VERSION AL551816.1 GI:12890126

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 1025)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

Location/Qualifiers

1. 1025

/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="CS0D1061918"
/clone_lib="L11_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com

BASE COUNT 282 a 255 c 268 g 207 t 13 others

ORIGIN

Query Match 22.6%; Score 212.2; DB 9; Length 1025;

Best Local Similarity 97.8%; Pred. No. 4.3e-31;

Matches 222; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 715 TCGGAGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGACGCGGCGGCTCGGA 775

Db 1 TCGGAGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGACGCGGCGGCTCGGA 60

QY 775 CCGGGCTCCGGGCTCGGTCAGCGGCGGCGGCGGCTCGGCGGCGGCGGCGG 834

Db 61 CCGGGCTCCGGGCTCGGTCAGCGGCGGCGGCGGCTCGGCGGCGGCGGCGG 120

QY 835 GGAAGTGTGTCTCTCTCGGCTCGGAGCGGCGGCGGCTCAGGCGGCGGCGGCGG 894

Db 121 GGAAGTGTGTCTCTCTCGGCTCGGAGCGGCGGCGGCTCAGGCGGCGGCGGCGG 180

QY 895 CGGCGAAGCGAGAGCGGACTCTGCGGCGGCGGCTCTTGCCCGCGGG 941

Db 181 CGGCGAAGCGAGAGCGGACTCTGCGGCGGCGGCTCTTGCCCGCGGG 227

RESULT 10

AL548886

LOCUS

DEFINITION

AL548886 L11_NFL006.PL2 Homo sapiens cDNA clone CS0D1043YK17.5

prime, mRNA sequence.

ACCESSION AL548886

VERSION AL548886.1 GI:12884334

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 638)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

Location/Qualifiers

1. 638

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1043YK17"

/clone_lib="L11_NFL006.PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with NotI and

cloned into the NotI and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax: (1) 301 610 8371

Db 121: GTGTCTCTGCTGGAGCCGCGAGCCGCGCTCAGGCGCGCGGGCGCGCGCGGCA 180
 Cy 902 CCAGAGACGACACTCTGCGCGCGCGGCTTTTGGCCGCGGG 941
 ||||||||||||||||||||||||||||||||||||||||
 Db 181 CCAGAGACGACACTCTGCGCGCGCGGCTTTGGCCGCGGG 220
 ||||||||||||||||||||||||||||||||||||||||

RESULT 14
 AL549158 740 bp mRNA linear EST 16-FEB-2001
 LOCUS AL549158 L11_NFL006_PL2 Homo sapiens cDNA clone CS001051P21 5
 DEFINITION prime, mRNA sequence.
 ACCESSION: AL549158
 VERSION AL549158.1 GI:12884867
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 740)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Pollayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91005 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
 Location/Qualifiers
 location 1..740
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS001051P21"
 /clone_lib="L11_NFL006_PL2"
 /tissue_type="placenta"
 /note="vector: pcwvsvprf 6; site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with NotI and
 cloned into the NotI and Eco RV sites of the pcwvsvprf 5
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : filiang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 177 a 195 c 228 g 138 t
 ORIGIN

Query Match 21.7%; Score 204.4; DB 9; Length 740;
 Best local Similarity 99.1%; Pred. No. 1.4e-29;
 Matches 216; Conservative 0; Mismatches 1; Gaps 1;

Cy 725 GCAGCGGACACTCTCTCGGCTCTCCCGGAGCGCGCGGCTCGAGCGCGGCTCC 784
 ||||||||||||||||||||||||||||||||||||||||
 Db 1 GCAGCGGACACTCTCTCGGCTCTCCCGGAGCGCGCGGCTCGAGCGCGGCTCC 60
 ||||||||||||||||||||||||||||||||||||||||

Cy 785 GGGGCTCGGGTCAGCGGCCAGC-GGGGGCTTGGCGGAGGATTACCGGGGAAGTGGT 843
 ||||||||||||||||||||||||||||||||||||||||
 Db 6 GGGGCTCGGGTCCAGCGGCCAGCGGGGGCCCTTGGCGGAGGATTACCGGGGAAGTGGT 120
 ||||||||||||||||||||||||||||||||||||||||

Cy 941 TGTCCTCTGGTGTAGAGCGCGGAGACGGGCGCTCAGAGCGCGGGCGCGCGCGGCAACG 903
 ||||||||||||||||||||||||||||||||||||||||
 Db 121 TGTCCTCTGGTGTAGAGCGCGGAGACGGGCGCTCAGAGCGCGGGCGCGCGCGGCAACG 180
 ||||||||||||||||||||||||||||||||||||||||

Cy 901 AGAGGACGACACTCTGCGCGCGGCTTTGGCCGCGGG 941
 ||||||||||||||||||||||||||||||||||||||||
 Db 181 AGAGGACGACACTCTGCGCGCGGCTTTGGCCGCGGG 218
 ||||||||||||||||||||||||||||||||||||||||

RESULT 15
 AL540382 857 bp mRNA linear EST 16-FEB-2001
 LOCUS AL540382 L11_F1002_P11 Homo sapiens cDNA clone CS0E001YK21 5 prim
 DEFINITION mRNA sequence.

```

ACCESSION AL540382
VERSION AL540382.1 GI:12870475
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope, Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LIL_FLO02_PL1"
/lab_host="DH10B"
/notes="Organ: Placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life technologies, a
division of Invitrogen 9800 Medical Center Drive, Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
flliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 219 a 226 c 250 g 170 t 2 others
ORIGIN
Query Match 21.5%; Score 202; DB 9; Length 867;
Best Local Similarity 98.2%; Pred. No. 3.9e-29;
Matches 224; Conservative 1; Mismatches 1; Indels 2; Gaps 2;
Oy 716 TCGAGCGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGAGCG-6CGGCGGCTCGG 774
Db 1 TCGAGCGCGCCAGCGGACACTCTCTCGGCTCTCTCCCGGAGCGGCGGCTCGG 50
Oy 775 AGCGGGCTCCGGGCTCGGTCGAGCGGCCAGCGG-GCGCTGGCGCGGAGATTACCG 833
Db 61 AGCGGCTCCGGGCTCGGTCGAGCGGCCAGCGGCGGCTGGCGGAGATTACCG 120
Oy 834 GCGAAGTGTTGTCTCTCGCTGGACCGCCAGACGGCGCTCAGGGCGCGGGCGCGG 893
Db 121 GCGAAGTGTTGTCTCTCGCTGGACCGCCAGACGGCGCTCAGGGCGCGGGCGCGG 180
Oy 894 GCGGCGACGAGAGCGGACTCTGGGCGCGGCTCTTTGGCCCGG 941
Db 181 GCGGCGACGAGAGCGGACTCTKGCGCGCGGCTCGTTGGCCCGG 228

```

Search completed: December 8, 2002, 12:49:39
 Job time : 1237.21 secs

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/930,548A
;; FILING DATE: 23-NOV-1992
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gordon, Alan M.
;; REGISTRATION NUMBER: 30,637
;; REFERENCE/DOCKET NUMBER: 31,298-01
;; TELEPHONE: 201-831-3244
;; TELEFAX: 201-831-3305
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4236 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..4068
;;
;; JS-07-930-548A-7

Query Match 25.3%; Score 76.2; DB 2; Length 4236;
Best Local Similarity 79.6%; Pred. No. 1.1e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 69 CTCGACTCTCTGGAAGTGGTCTCTGAGTATCTTTGACAAATCTACAGCACCA 128

DB 3194 CTCCTCCCTCTTGAAGTGGTGGCCCGGAGAACAAATTTTGACAGACTGTACACATCC 3253

QY 129 AGACGACGTGGTCTTACGAGTATGCTGTGGGAAATCTTCCTTAGGT 181

DB 3254 AGAGTGACGTGCTGTTGCTTTGCTGTGGGAAATATTTTCCTTAGGT 3306

RESULT 5
US-07-813-593-3
; Sequence 3, Application US/07/813593
; Patent No. 5185438
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/813,593
; FILING DATE: 19920415
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601

;; REFERENCE/DOCKET NUMBER: LEM-3-PPP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-645-1405
;; TELEFAX: 212-645-2054
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5406 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 208..4311
;; NAME/KEY: Rat_peptide
;; LOCATION: 208..4308
;;
;; US-07-813-593-3

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 69 CTCGACTCTCTGGAATGGATGGTCTCTGAGTATCTTTGACAAATCTACAGCACCA 128

DB 3395 CCGAGCTCCCTTTGAGTGGATGGCCCGGAGAACAAATTTTGACAGAGTATACACAAATTC 3454

QY 129 AGACGACGTGGTCTTACGAGTATGCTGTGGGAAATCTTCCTTAGGT 181

DB 3455 AGACGAGTGGTCTTACGAGTATGCTGTGGGAAATATTTTCCTTAGGT 3507

RESULT 6
US-07-977-451-5
; Sequence 5, Application US/07977451
; Patent No. 5270458
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 19921119
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
US-07-977-451-5

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTGCACTTCTCTGAATGAGTGGCTCCGTAATCTATCTTTGACAAATCTACAGCACC 128
DB 3395 CCCGACTCCCTTGAAGTGGGCGCCGAAACCAATTTTCACAGAGTATACAAATTC 3454
QY 129 AGAGCCACGTGGTCTTACGAGATATCTGTGGGAATCTTCTCTTAGGT 181
DB 3455 AGAGCGATGTGGTCTTCTGGTGTGTCTGTGGGAATATTTTCTTAGGT 3507

RESULT 7
US-07-946-507-3
Sequence 3, Application US/07946507
Patent No. 5283354
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,507
FILING DATE: 19920917
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
NAME/KEY: mat_peptide
LOCATION: 208..4308
US-07-946-507-3

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTGCACTTCTCTGAATGAGTGGCTCCGTAATCTATCTTTGACAAATCTACAGCACC 128
DB 3395 CCCGACTCCCTTGAAGTGGGCGCCGAAACCAATTTTCACAGAGTATACAAATTC 3454
QY 129 AGAGCCACGTGGTCTTACGAGATATCTGTGGGAATCTTCTCTTAGGT 181
DB 3455 AGAGCGATGTGGTCTTCTGGTGTGTCTGTGGGAATATTTTCTTAGGT 3507

RESULT 8
US-08-252-517-5
Sequence 5, Application US/08252517
Patent No. 5548065
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/252,517
FILING DATE: 31-OCT-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: TW 81102961
FILING DATE: 19-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 265..4308
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 208..264
US-08-252-517-5

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTCGACTTCCTCGAATGATGCTGCTGAATCTATCTTTGACAAATCTACAGCACCA 128
Db 3395 CCGGACTCCCTTGAAGTGGATGGCCCGGAACCATTTTGACAGATATACACATTC 3454
QY 129 AGAGCGAGGTGGTCTTACGAGGATTTGCTGGGAAATCTTCTTCCTTAGGT 181
Db 3455 AGAGCGAGGTGGTCTTCCGCTGCTGGGAAATATTTTCCTTAGGT 3507
RESULT 9
US-07-906-397A-5

Sequence 5, Application US/07906397A
Patent No. 5621090
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
STREET: 180 VARICK STREET
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,397A
FILING DATE: 19920626
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-PPTPPP
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5406 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 208..4311
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 208..4308
US-07-906-397A-5

Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTCGACTTCCTCGAATGATGCTGCTGAATCTATCTTTGACAAATCTACAGCACCA 128
Db 3395 CCGGACTCCCTTGAAGTGGATGGCCCGGAACCATTTTGACAGATATACACATTC 3454
QY 129 AGAGCGAGGTGGTCTTACGAGGATTTGCTGGGAAATCTTCTTCCTTAGGT 181
Db 3455 AGAGCGAGGTGGTCTTCCGCTGCTGGGAAATATTTTCCTTAGGT 3507

RESULT 10
US-08-601-891-5
; Sequence 5, Application US/08601891
; Patent No. 5747651
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varlick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/601.891
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5406 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 208...4311
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: 265...4308
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 208...264
; US-08-601-891-5
Query Match 24.8%; Score 74.6; DB 1; Length 5406;
Best Local Similarity 78.8%; Pred. No. 4.1e-14;
Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 69 CTCGACTTCCTCTGAAATGAGTGGCTCTGATCTATCTTTGACAAAATCTACAGACCA 128
DB 3395 CCGGACTCCCTTTGAAATGAGTGGCTCTGATCTATCTTTGACAAAATCTACAGACCA 128
QY 129 AGAGCGAGTGGTGGCTTTACGGAGTATGCTGTGGGAAATCTTCCTTACG 181
DB 3455 AGAGCGAGTGGTGGCTTTACGGAGTATGCTGTGGGAAATCTTCCTTACG 3507
RESULT 11
US-09-021-324-5
; Sequence 5, Application: US/09021324
; Patent No. 5912133
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varlick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021.324
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 04:25:08 ; Search time 330.95 Seconds
(without alignments)
2048.198 Million cell updates/sec

Title: US-09-778-900A-3_COPY_200_500
Perfect score: 301
Sequence: 1 ggacactcgggttgatgtg.....taacaaaaaacctttaaagt 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
tal number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	301	100.0	726	22 AAD19454	Human flt-1 gene D
2	114.4	38.0	4017	24 ABL91752	Human polynucleoti
3	112.8	37.5	7680	23 AAS70286	DNA encoding novel
4	76.2	25.3	4044	24 ABL91751	Human polynucleoti
5	76.2	25.3	4071	13 AAQ28272	A novel type III R
6	76.2	25.3	4071	20 AAV99829	Human receptor tyr
7	76.2	25.3	4071	24 ABL91754	Human polynucleoti
8	76.2	25.3	4225	22 AAF83308	Human VEGFR-2 enco
9	76.2	25.3	4236	19 AAV34763	Human KDR genomic

10	74.6	24.8	5404	17 AAT38735	Murine foetal live
11	74.6	24.8	5404	20 AAX77516	Murine flk-1 cDNA.
12	74.6	24.8	5406	13 AAQ29957	flk-1 cDNA sequenc
13	74.6	24.8	5406	14 AAQ53504	Murine flk-1 cDNA.
14	74.6	24.8	5406	14 AAQ35251	Human flk-1 coding
15	74.6	24.8	5406	14 AAQ40316	Murine flk-1 cDNA.
16	74.6	24.8	5406	16 AAQ81014	Flk1 receptor prot
17	74.6	24.8	5406	16 AAQ79070	Mouse flk-1 cDNA.
18	74.6	24.8	5406	18 AAT72119	Murine flk-1 recep
19	74.6	24.8	5470	15 AAQ64049	Sequence of murine
20	67.2	22.3	2949	19 AAV39041	Human receptor typ
21	67.2	22.3	2958	19 AAV39040	Human receptor typ
22	67.2	22.3	2958	19 AAV39042	Human receptor typ
23	67.2	22.3	2978	19 AAV39038	Human receptor typ
24	67.2	22.3	2982	19 AAV39039	Human receptor typ
25	67.2	22.3	3120	14 AAQ49756	PK gene lptk25.
26	67.2	22.3	3120	16 AAT03096	Protein tyrosine-k
27	67.2	22.3	3475	16 AAT00802	Human Flk2/flt3 ty
28	67.2	22.3	3476	16 AAQ91536	Human STK-1 cDNA.
29	67.2	22.3	3501	14 AAQ53503	Human flk-2 cDNA.
30	67.2	22.3	3501	14 AAQ35250	Human flk-2 coding
31	67.2	22.3	3501	14 AAQ40315	Human flk-2 cDNA.
32	67.2	22.3	3501	16 AAQ81013	Flk2 receptor prot
33	67.2	22.3	3501	16 AAQ79069	Human flk-2 cDNA.
34	67.2	22.3	3501	17 AAT38734	Murine foetal live
35	67.2	22.3	3501	18 AAT72117	Human flk-2 recept
36	67.2	22.3	3501	20 AAX77515	Human flk-2 cDNA.
37	65.6	21.8	3897	24 ABL91753	Human polynucleoti
38	65.6	21.8	4111	21 AAC62405	Human Flt4/VEGFR-3
39	65.6	21.8	4111	22 AAC68952	Human Flt4/VEGFR-3
40	65.6	21.8	4195	17 AAT12068	Flt4 receptor tyro
41	65.6	21.8	4195	21 AAA37815	Human Flt4 recepto
42	65.6	21.8	4195	21 AA252333	Human tyrosine kin
43	65.6	21.8	4425	16 AAT03090	Protein tyrosine-k
44	65.6	21.8	4450	21 AAC62210	Nucleotide sequenc
45	65.6	21.8	4795	21 AAA37816	Human Flt4 recepto

ALIGNMENTS

RESULT 1
AAD19454
ID AAD19454 standard: DNA; 726 BP.
XX
AC AAD19454;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human flt-1 gene DNA fragment #3.
XX
KW Human flt-1 gene; cytostatic; haemostatic; gene mapping; gynaecological;
KW optalmological; rheumatoid arthritis; endometriosis; angiogenic disease;
KW diabetic retinopathy; psoriasis; VEGFR-1; drug therapy; pharmacogenetic;
KW vascular endothelial growth factor; single nucleotide polymorphism; SNP;
KW cancer; radicant; ds.
XX
OS Homo sapiens.
XX
FH Key
FT intron
FT exon
FT intron
FT variation
FT replace (454, A)
FT /tag= a
FT /tag= b
FT /tag= c
FT /tag= d
FT /standard_name= *Single nucleotide polymorphism (SNP)*
PN EP11:0123-A2.
XX
PD 05-SEP-2001.

XX PF 20-FEB-2001; 2001EP-0301489.
XX PR 24-FEB-2000; 2000GB-0004232.
XX PA (ASTR) ASTRAZENECA AB.
XX PI Smith JC;
XX DR WPI; 2001-608199/70.
XX PT Diagnosing single nucleotide polymorphism in human flt-1 gene, for
PI assessing predisposition or susceptibility to diseases mediated by an
PI flt-1 ligand, comprises determining the nucleic acid sequence at one or
PI more point positions -
XX PS Claim 8; Page 23; 30pp; English.
XX CC The invention relates to novel sequence and polymorphisms in the human
CC flt-1 gene. Flt-1 is one of the two receptors for vascular endothelial
CC growth factor (VEGF-1). The invention also relates to a method for
CC diagnosing one or more single nucleotide polymorphism (SNP) in human
CC flt-1 gene. The method is useful for assessing the predisposition and/or
CC susceptibility of an individual to diseases mediated by an flt-1 ligand,
CC to recognise individuals who are particularly at risk from developing
CC these conditions, and in the development of new drug therapies that
CC selectively target one or more allelic variants of the flt-1 gene. Flt-1
CC polymorphisms are used in the diagnosis and treatment of cancer and
CC angiogenic diseases such as diabetic retinopathies, psoriasis, rheumatoid
CC arthritis and endometriosis. Polymorphisms are also used in mapping the
CC human genome and to elucidate the genetic component of the diseases, as
CC genetic markers for this region in linkage studies. The nucleic acids are
CC useful as components in databases useful in sequence identity, genome
CC mapping, pharmacogenetics and other search analyses. The flt-1 antagonist
CC is useful in the preparation of a medicament for treating VEGF-mediated
CC diseases in a human diagnosed as having a single nucleotide polymorphism
CC at one or more of the defined positions. The present sequence is human
CC flt-1 gene DNA fragment related to the invention. This DNA fragment
CC contains exon 24 and adjacent intron sequences.
XX SQ Sequence 726 BP; 199 A; 165 C; 154 G; 207 T; 1 other;
Query Match 100.0%; Score 301; DB 22; Length 726;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGACATCGGGTGAATGCTTTGTTTTCATCCCTTCCTCAAGGCCCTTGCCT 60
DB 200 GGACATCGGGTGAATGCTTTGTTTTCATCCCTTCCTCAAGGCCCTTGCCT 259
61 CTTCTAGACTGCTCTCTCTGAATGATGGTCTCTGAATCTATCTTIGACAAATCTA 120
DB 260 CTTCTAGACTGCTCTCTCTGAATGATGGTCTCTGAATCTATCTTIGACAAATCTA 319
QY 121 CAGCACCAGAGCGGTGCTCTTACGGAGTATTGCTGTGGGAATCTTCCTTAGG 180
DB 320 CAGCACCAGAGCGGTGCTCTTACGGAGTATTGCTGTGGGAATCTTCCTTAGG 379
QY 181 TAAATTTGGGAGGAGGAATCAACACAGCCAGAAATAATGTCTGCATCTTCTCTG 240
DB 380 TAAATTTGGGAGGAGGAATCAACACAGCCAGAAATAATGTCTGCATCTTCTCTG 439
QY 241 AATGTCCTTTGGTGGACAGCCTTTAGATTAGAACCTACTGTACAAAAAACTCTTAAG 300
DB 440 AATGTCCTTTGGTGGACAGCCTTTAGATTAGAACCTACTGTACAAAAAACTCTTAAG 499
QY 301 T 301
DB 500 T 500
RESULT 2
ABL91752

ID ABL91752 standard; cDNA; 4017 BP.
XX ABL91752;
XX AC
XX XX
XX DT 28-MAY-2002 (first entry)
XX XX
XX DE Human polynucleotide SEQ ID NO 95.
XX XX
XX KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
XX KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
XX KW cytosolic; virucide; protozoicide; antibacterial; ds.
XX CS Homo sapiens.
XX XX
XX FN DE10100586-Cl.
XX XX
XX FD 11-APR-2002.
XX XX
XX PF 09-JAN-2001; 2001DE-1000586.
XX XX
XX PR 09-JAN-2001; 2001DE-1000586.
XX PA (RIBO-) RIBOPHARMA AG.
XX PI Kreutzer R, Lillmer S, Rost S, Hadwiger P;
XX DR WPI; 2002-270454/32.
XX XX
XX PI Inhibiting gene expression in cells, useful for e.g. treating tumors,
PI by introducing double-stranded complementary oligonucleotide having unpaired
PI terminal bases
XX PS Claim 13; Page 68-69; 104pp; German.
XX CC The invention relates to a method for inhibiting expression of a target
CC gene (ABL91658-AB91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a
CC single-stranded segment of 1-4 nt. The method provides
CC oligoribonucleotides for antisense inhibition of gene expression useful
CC e.g. for treating tumours but the oligoribonucleotides may also be
CC directed against genes present in pathogens (e.g. Plasmodium or
CC viruses/viroids, pathogenic on humans, animals or plants) or against
CC cytokine, ID, developmental or prion genes. The method provides more
CC effective inhibition of gene expression than use of known
CC oligonucleotides, probably because the unpaired overhang increases
CC stability and thus intracellular concentration.
XX SQ Sequence 4017 BP; 1236 A; 915 C; 903 G; 963 T; 0 other;
Query Match 38.0%; Score 114.4; DB 24; Length 4017;
Best Local Similarity 99.1%; Pred. No. 2.5e-26;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 66 AGACTCCACTTCCTTGAATGATGGTCTCTGACTATCTTTCACAAATCTACAGCA 125
DB 3173 ATACTCGACTTCCTTGAATGATGGTCTCTGACTATCTTTCACAAATCTACAGCA 3232
QY 126 CCAAGACGCGGTGGTCTTACGGAGTATTGCTGTGGGAATCTTCTCCTTAGGT 181
DB 3233 CCAGACGCGCGGTGGTCTTACGGAGTATTGCTGTGGGAATCTTCTCCTTAGGT 3288
RESULT 3
AAS70286
ID AAS70286 standard; cDNA; 7680 BP.
XX AAS70286;
XX AC
XX XX
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #6090.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW Homo sapiens.
OS WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG06099.
XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity
XX Claim 1; SEQ ID No 6090; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 7680 BP; 2279 A; 1661 C; 1739 G; 2001 T; 0 other;
Query Match 37.5%; Score 112.8; DB 23; Length 7680;
Best Local Similarity 98.3%; Pred. No. 1.1e-25;
Matches 114; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 66 AGACTGACCTCTCTGAATGGATGGCTCTCTGAATCTATCTTTGACAAATCTACAGCA 125
DB 3422 ATACTGACCTCTCTGAATGGATGGCTCTCTGAATCTATCTTTGACAAATCTACAGCA 3481
QY 126 CCAGACGACGTGGCTTACGAGTATTCGTGGGAAATCTCTCTTAGGT 181
DB 3482 CCAAGACGACGTGGCTTACGAGTATTCGTGGGAAATCTCTCTTAGGT 3537
RESULT 4
ABL91751
ID ABL91751 standard; DNA; 4044 BP.
XX ABL91751;
XX 28-MAY-2002 (first entry)
DT

XX Human polynucleotide SEQ ID NO 94.
DE Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen; Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide; cytostatic; virucide; protozoacide; antibacterial; ds.
XX Homo sapiens.
XX DE10100586-C1.
XX 11-APR-2002.
XX 09-JAN-2001; 2001DE-1000586.
XX 09-JAN-2001; 2001DE-1000586.
XX (RIBO-) RIBOPHARMA A3.
XX K-reutner R, Limmer S, Rost S, Hadwiger P;
XX WPI; 2002-270454/32.
XX Inhibiting gene expression in cells, useful for e.g. treating tumors, by introducing double-stranded complementary oligoRNA having unpaired terminal bases
XX Claim 13; Page 67-68; 104pp; German.
XX The invention relates to a method for inhibiting expression of a target gene (AAS16158-ABL51797) in a cell by introducing at least one oligoribonucleotide that has a double-stranded structure consisting of at most 49 sequential nucleotide pairs, with at least part of one strand complementary with the target gene and has at least one end a single-stranded segment of 1-4 nt. The method provides oligoribonucleotides for antisense inhibition of gene expression useful e.g. for treating tumors but the oligoribonucleotides may also be directed against genes present in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic on humans, animals or plants) or against cytokine, Id, developmental or prion genes. The method provides more effective inhibition of gene expression than use of known oligonucleotides, probably because the unpaired overhang increases stability and thus intracellular concentration.
XX Sequence 4044 BP; 1163 A; 885 C; 1021 G; 975 T; 0 other;
Query Match 25.3%; Score 76.2; DB 24; Length 4044;
Best Local Similarity 79.6%; Pred. No. 5e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 69 CTCGACTTCTCTGAATGGATGGCTCTCTGAATCTATCTTTGACAAATCTACAGCA 128
DB 3194 C1CGCTCTCTCTGAATGGATGGCTCTCTGAATCTATCTTTGACAGAGTACACATCC 3253
QY 129 AGACGACGCTGGCTCTACGAGTATTCGTGGGAAATCTCTCTTAGGT 181
DB 3254 AGAGTACGCTGGCTCTCTTTGGTTTGGTGGGAAATCTCTCTTAGGT 3306
RESULT 5
AAQ28272
ID AAQ28272 standard; cDNA; 4071 BP.
XX AAQ28272;
XX 15-FEB-1993 (first entry)
DT
XX A novel type III RIK gene - the KDR gene.
XX Receptor tyrosine kinase; vascular endothelial cell growth factors; cancer; tumour; diagnosing; monitoring; ss.
XX Homo sapiens.
OS

```
XX PH Key Location/Qualifiers
XX FT CDS 1..4068
XX FT /*Tag= a
XX PN W09214748-A.
XX PD 03-SEP-1992.
XX PR 20-FEB-1992; 92WO-US01300.
XX PR 22-FEB-1991; 91US-0657236.
XX PR (AMCY ) AMERICAN CYANAMID CO.
XX PI Carrión ME, Terman BI;
XX NPI: 1992-316117/38.
XX P-PSDB: AAR26999.
XX DNA encoding type III receptor tyrosine kinase - useful for
  diagnosing the onset of cancer
XX Claim 3; Fig 7; 101pp; English.
XX This sequence represents a novel type III receptor tyrosine kinase
  gene. A labelled EcoRI-BamHI DNA segment derived from clone Bf111081.8
  was used as a probe to rescreen a human endothelial cDNA library
  (HLI0246) for 5' full length DNA segments of the gene from which the
  insert portion of Bf111081.8 is derived. A synthetic probe designed
  from nucleotides 3297-3325 of Bf111081.8 is then used to isolate
  more 3' full length clones. One of the clones, designated Bf111200.2
  is cloned into pBluescript KS and the synthetic oligonucleotide
  JGACGCCG ATC GAG cloned, which contains the initial sequence Met-
  Glu, the first two amino acids encoded by the KDR gene, forming
  Bf11140, which is then purified on a CsCl density gradient. This
  was sequenced, together with Bf111081.8, and Bf11129.5 to comprise
  the entire ORF of 4,068 nucleotides of the KDR gene.
XX Sequence 4071 BP; 1172 A; 894 C; 1024 G; 981 T; 0 other;
Query Match 25.3%; Score 76.2; DB 13; Length 4071;
Best Local Similarity 79.6%; Pred. No. 5e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 69 CTCGACTTCTCTGAAATGGATGGCTCTGATCTATCTTTCACAAATCTACAGCACCA 128
DB 3194 CTCGCTCCCTTTGAAATGGATGGCTCTGATCTATCTTTCACAAATCTACAGCACCA 128
DB 129 AGAGCGACGTGGTCTTACGAGATATGCTGTGGGAAATCTTCTCTTAGGT 181
DB 3254 AGAGTACGCTGGCTCTTGGTGTGTTGCTGTGGGAAATATTTTCTTAGGT 3306
RESULT 6
AAV99829
ID AAV99829 standard; cDNA; 4071 BP.
XX AC AAV99829;
XX DT 12-APR-1999 (first entry)
XX Human receptor tyrosine kinase KDR cDNA.
XX KDR; receptor tyrosine kinase; human; signal transduction;
XX mitogen; neoplasia; angiogenesis; diabetes; retinopathy;
XX breast cancer; brain cancer; inflammation; rheumatoid arthritis;
XX psoriasis; contact dermatitis; hypersensitivity; antagonist;
XX gene therapy; ds.
XX Homo sapiens.
XX OS
XX PN W09858053-A1.
```

```
XX PD 23-DEC-1998.
XX PR 17-JUN-1998; 98WO-US12569.
XX PR 18-JUN-1997; 97US-0050962.
XX PR (MERI ) MERCK & CO INC.
XX Kendall RL, Mao X, Tebben A, Thomas KA;
XX NPI: 1999-095333/08.
XX P-PSDB: AAW80997.
XX Human receptor tyrosine kinase protein, KDR - useful e.g. to screen
  for antagonists useful to treat diseases involving neovascularization
  e.g. diabetic retinal vascularization, cancers
XX Claim 1; Fig 1a; 69pp; English.
XX This nucleotide sequence encodes a novel receptor tyrosine kinase,
  termed KDR (see AAW80997), that is expressed on human endothelial
  cells. KDR is activated by vascular endothelial growth factor and
  mediates a mitogenic signal. It is implicated in clinical
  neovascularization. KDR cDNA was isolated from a human umbilical vein
  endothelial cell lambda phage cDNA library using a 576-bp DNA probe
  that had been generated by PCR (see also AAW9850-53). The predicted
  protein product has amino acid differences from the previously
  published KDR sequence at positions 498 (Ala to Glu), 772 (Thr to
  Ala), 787 (Gly to Arg), 835 (Asn to Lys), 848 (Glu to Val) and 1347
  (Thr to Ser), producing a protein predicted by computer modeling to
  have higher activity and functionality. The invention also relates
  to recombinant vectors and recombinant hosts which contain a DNA
  fragment encoding human KDR, a DNA fragment encoding the
  intracellular portion of KDR with or without a membrane anchor
  sequence, purified forms of associated human KDR, and human mutant
  forms of KDR. KDR, fusion proteins or fragments can be used in
  assays to identify antagonists and agonists of human KDR (claimed).
  Antagonists of KDR useful for treating diseases involving
  neovascularization e.g. diabetic retinal vascularization, cancers
  (e.g. brain, breast, etc.) and forms of inflammation e.g. rheumatoid
  arthritis, psoriasis, contact dermatitis and hypersensitivity
  reactions. The polynucleotides are useful to screen for KDR
  antagonists/agonists and for gene therapy (e.g. by introducing a
  gene portion encoding a KDR protein containing functional ligand
  binding and membrane anchoring moieties but not tyrosine kinase
  activity). They are also useful to measure levels of human KDR.
XX Sequence 4071 BP; 1169 A; 894 C; 1027 G; 981 T; 0 other;
Query Match 25.3%; Score 76.2; DB 20; Length 4071;
Best Local Similarity 79.6%; Pred. No. 5e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 69 CTCGACTTCTCTGAAATGGATGGCTCTGATCTATCTTTCACAAATCTACAGCACCA 128
DB 3194 CTCGCTCCCTTTGAAATGGATGGCTCTGATCTATCTTTCACAAATCTACAGCACCA 128
QY 129 AGAGCGACGTGGTCTTACGAGATATGCTGTGGGAAATCTTCTCTTAGGT 181
DB 3254 AGAGTACGCTGGCTCTTGGTGTGTTGCTGTGGGAAATATTTTCTTAGGT 3306
RESULT 7
ABL91754
ID ABL91754 standard; DNA; 4071 BP.
XX AC ABL91754;
XX DT 28-MAY-2002 (first entry)
XX Human polynucleotide SEQ ID NO 97.
```

```
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS Homo sapiens.
XX
PN DE10100586-C1.
XX
PD 11-APR-2002.
XX
PF 09-JAN-2001; 2001DE-1000586.
XX
PR 09-JAN-2001; 2001DE-1000586.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX WPI; 2002-270454/32.
XX
PS Claim 13; Page 71-72; 104pp; German.
XX
CC The invention relates to a method for inhibiting expression of a target
CC gene (ASL91658-ASL91797) in a cell by introducing at least one
CC oligoribonucleotide that has a double-stranded structure consisting of at
CC most 49 sequential nucleotide pairs, with at least part of one strand
CC complementary with the target gene and has at least one end a
CC single-stranded segment of 1-4 nt. The method provides
CC oligoribonucleotides for antisense inhibition of gene expression useful
CC e.g. for treating tumours but the oligoribonucleotides may also be
CC directed against genes present in pathogens (e.g. Plasmodium or
CC viruses/viroids, pathogenic on humans, animals or plants) or against
CC cytokine, Id, developmental or prion genes. The method provides more
CC effective inhibition of gene expression than use of known
CC oligonucleotides, probably because the unpaired overhang increases
CC stability and thus intracellular concentration.
XX
SQ Sequence 4071 BP; 1169 A; 894 C; 1025 G; 983 T; 0 other;

Query Match 25.3%; Score 76.2; DB 24; Length 4071;
Best Local Similarity 79.6%; Pred. No. 5e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

69 CTCGACTTCTCTGAAATGGATGGCTCTCGAATCTATCTTTGACAAATCTACAGCACCA 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3194 CTCGCTCTCTCTGAAATGGATGGCTCTCGAATCTATCTTTGACAAATCTACAGCACCA 3253

QY 129 AGAGCGACGTGGTGGCTTACCGAGTATTGCTGGGAAATCTTCTCTTAGGT 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3254 AGAGTGACGTGGTGGCTTACCGAGTATTGCTGGGAAATCTTCTCTTAGGT 3306

RESULT 8
AAF83308
ID AAF83308 standard; DNA; 4225 BP.
XX
AC AAF83308;
XX
DI 09-JUL-2001 (first entry)
XX
DE Human VEGFR-2 encoding DNA.
XX
KW Receptor protein; vascular endothelial growth factor receptor-2;
KW VEGFR-2; neuropilin-1; NP-1; co-receptor; human; angiogenic; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 70..4140
```

```
FI /*tag= a
FI /product= "VEGFR-2"
FI /misc_feature 71..2350
FI /*tag= b
FI /note= "extracellular domain coding fragment (AAF83310)."
XX
PN WC200131346-A2.
XX
PD 03-MAY-2001.
XX
PF 26-OCT-2000; 2000WC-US29579.
XX
PR 28-OCT-1999; 99US-0162367.
XX
PA (PROC ) PROCTER & GAMBLE CO.
XX
PI Rosenbaum JS, Whitaker GB, Limberg BJ;
XX WPI; 2001-308686/32.
XX P-PSDB; AAB562475.
XX
CC Determining compounds which bind to a complex comprising vascular
CC endothelial growth factor receptor-2 and Neuropilin-1 to provide
CC superior pro- and anti-angiogenic agents
XX
PS Example 1; Page 32-39; 82pp; English.
XX
CC The invention relates to determining whether a compound is capable of
CC binding to a receptor protein complex comprising a vascular endothelial
CC growth factor receptor-2 (VEGFR-2) receptor protein and a neuropilin-1
CC (NP-1) receptor protein. One method comprises introducing a sample
CC comprising the compound to the receptor protein and allowing the
CC compound to bind to the complex. Signaling through VEGFR-2 is enhanced
CC in the presence of the NP-1 co-receptor. The methods of the invention
CC can be used for identifying novel pro- and anti-angiogenic compounds.
CC The present sequence represents the DNA encoding a human VEGFR-2.
XX
SQ Sequence 4225 BP; 1304 A; 940 C; 1067 G; 1014 T; 0 other;

Query Match 25.3%; Score 76.2; DB 22; Length 4225;
Best Local Similarity 79.6%; Pred. No. 5.1e-14;
Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 69 CTCGACTTCTCTGAAATGGATGGCTCTCGAATCTATCTTTGACAAATCTACAGCACCA 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3263 CTCGCTCTCTCTGAAATGGATGGCTCTCGAATCTATCTTTGACAGAGTGTACACAATCC 3322

QY 129 AGAGCGACGTGGTGGCTTACCGAGTATTGCTGGGAAATCTTCTCTTAGGT 181
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3323 AGAGTGACGTGGTGGCTTACCGAGTATTGCTGGGAAATCTTCTCTTAGGT 3375

RESULT 9
AAV34763
ID AAV34763 standard; DNA; 4236 BP.
XX
AC AAV34763;
XX
DI 27-AUG-1998 (first entry)
XX
DE Human KDR genomic DNA.
XX
KW Kinase insert domain, containing receptor; KDR; screening; inhibitor;
KW vascular endothelial growth factor; VEGF; angiogenesis; treatment;
KW cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..4071
FT /product= KDR
FT /note= "kinase insert domain containing receptor"
FT
```

```

XX PN US5766860-A.
XX PD 16-JUN-1998.
XX PF 25-FEB-1997; 97US-0810116.
XX PR 23-NOV-1992; 92US-0930548.
XX PR 25-FEB-1997; 97US-0810116.
XX PA (AMCY ) AMERICAN CYANAMID CO.
XX PA Carrión ME, Terman BI;
XX PF 1998-361682/31.
XX PR P-PSDB; AAW59275.
XX PT Screening assay for vascular endothelial cell growth factor
XX PT antagonists - using recombinant cells expressing receptor protein
XX PT Claim 1: Flg 7A-X; 5lpp; English.
XX
XX This sequence encodes a novel human growth factor receptor, kinase
XX insert domain containing receptor or KDR. This receptor is capable of
XX binding to the vascular endothelial cell growth factor, VEGF and is
XX used in a screening assay which identifies compounds that inhibit VEGF
XX action on KDR. Such compounds which inhibit binding of VEGF to the KDR
XX may inhibit angiogenesis and thus be useful for treating cancer.
XX
XX Sequence 4236 BP; 1216 A; 938 C; 1062 G; 1020 T; 0 other;
XX
XX Query Match 25.3%; Score 76.2; DB 19; Length 4236;
XX Best Local Similarity 79.6%; Pred. No. 5.1e-14;
XX Matches 90; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
XX
XX QY 69 CTCGACTTCTCGAATGATGGCTCCCTGAAATCTATCTTTCGAAATCTACACACCA 128
XX DB 3194 CTCGCTCCCTTGAATGATGGCTCCGAGAAATCTTTCGAGAGTACACATCC 3253
XX
XX QY 129 AGAGCAGCTGTGGCTTACGAGATTCCTGTGGAAATCTTCTCTTAGGT 181
XX DB 3254 AGAGTACGCTGTGGCTTTCGAGTTCCTGTGGAAATCTTCTCTTAGGT 3306
XX
XX RESULT 10
XX AAT38735
XX ID AAT38735 standard; cDNA; 5404 BP.
XX SC AAT38735;
XX
XX 11-DEC-1996 (first entry)
XX
XX DE Murine foetal liver kinase 1 cDNA.
XX
XX KW Murine; foetal liver kinase 1; flk-1; protein tyrosine kinase;
XX KW monoclonal; antibody; extracellular domain; receptor assay;
XX KW haematopoietic stem cell; ligand; stimulation; proliferation;
XX KW differentiation; treatment; anaemia; bone marrow damage;
XX KW cancer chemotherapy; radiation; ds.
XX
XX OS Mus musculus.
XX
XX FH Key Location/Qualifiers
XX FT CDS 208..4311
XX FT sig_peptide 208..264
XX FT mat_peptide 265..4308
XX FT /tag= a
XX FT /tag= b
XX FT /tag= c
XX
XX PN US5548065-A.
XX PD 20-AUG-1996.

```

```

XX PF 02-APR-1991; 91US-0679666.
XX PR 15-NOV-1992; 92US-0977451.
XX PR 02-APR-1991; 91US-0679666.
XX PR 28-JUN-1991; 91US-0726913.
XX PR 15-NOV-1991; 91US-0793065.
XX PR 24-DEC-1991; 91US-0813593.
XX PR 26-JUN-1992; 92US-0806397.
XX PR 12-NOV-1992; 92US-0975049.
XX PR 30-APR-1993; 93US-0055269.
XX PR 31-OCT-1994; 94US-0252517.
XX (UYPR-) UNIV PRINCELOON.
XX PA Lemischka IR;
XX PF 1996-392678/39.
XX PR P-PSDB; AAT38735.
XX PT Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays,
XX PT for isolating haematopoietic stem cells expressing receptor and for
XX PT obtaining ligands
XX
XX Disclosure: Columns 51-62; 50pp; English.
XX
XX The present sequence encodes murine foetal liver kinase 1 (flk-1),
XX a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
XX raised against the extracellular portion of flk-1 can be used to
XX assay for flk receptors on the surface of haematopoietic stem
XX cells, and to isolate positive cells. The antibodies can also
XX be used as, or to obtain ligands, which stimulate the proliferation
XX and/or differentiation of stem cells. The ligands can be used, e.g.
XX for treating anaemia, or bone marrow damage resulting from cancer
XX chemotherapy, or radiation.
XX
XX Sequence 5404 BP; 1411 A; 1297 C; 1423 G; 1273 T; 0 other;
XX
XX Query Match 24.8%; Score 74.6; DB 17; Length 5404;
XX Best Local Similarity 78.8%; Pred. No. 1.8e-13;
XX Matches 89; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
XX
XX QY 69 CTCGACTTCTCGAATGATGGCTCCCTGAAATCTATCTTTCGAAATCTACACACCA 128
XX DB 3395 CCGAGCTCCCTTGAATGATGGCTCCGAGAAATCTTTCGAGAGTACACATTC 3454
XX
XX QY 129 AGAGCAGCTGTGGCTTACGAGATTCCTGTGGAAATCTTCTCTTAGGT 181
XX DB 3455 AGAGCAGCTGTGGCTTTCGAGTTCCTGTGGAAATCTTCTCTTAGGT 3507
XX
XX RESULT 11
XX AAX77516
XX ID AAX77516 standard; cDNA; 5404 BP.
XX AC AAX77516;
XX DT 05-AUG-1999 (first entry)
XX DE Murine flk-1 cDNA.
XX
XX KW Murine; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
XX KW monoclonal; polyclonal; antibody; tyrosine kinase; ds.
XX
XX OS Mus sp.
XX
XX FH Key Location/Qualifiers
XX FT CDS 208..4311
XX FT /tag= a
XX FT /product= "flk-1"
XX
XX PN US5912133-A.

```


Db 3267 GAAGTGTATCCACAGGACCTGGCAGCAGGAACATTCCTATCGGAGAAGATGSGT 3326
QY 1458 GAAGATTGTGATTTTGGCCCTTCCCGGGATATTTATAGAACCCCGATTAATGTCAGAAA 1517
Db 3327 TAAGATCTGTGATTCGGCTTGGCCCGGGACATTTAAGACCCCGATTAATGTCAGAAA 3386
QY 1518 AGGAGATACTCGACTTCCTCTGAAATGATGGCTCCCGAATCTATCTTTGACAAAATCTA 1577
Db 3387 AGGAGATCCCGACTCCCTTTCAAGTGTGATGGCCCGGAACCATTTTGCACAGTATA 3446
QY 1578 CAGCACCAAGAGCGAGTGTGGCTTTAGCGAGTATTCCTGTGGGAATCTTCTCTTAGG 1637
Db 3447 CACAATTCAGAGCGATGTGTGGCTTTAGCGAGTATTCCTGTGGGAATTTTCTTAGG 3506
QY 1638 TGGGTCTCCATACCAGGAGTACAAATGGATGAGGACTTTTGCAGTCCGCTGAGGGAAGG 1697
Db 3507 TCCCTCCCATCCCTGGGGTCAAGATTGATGAAGATTTTGTAGAGATTGAANGAGG 3566
QY 1698 CATGAGTATGAGAGCTCCTGTAGTACTCTCTGAAATCTATCAGATCATCTGGACTG 1757
Db 3567 AACTAGATGGGGTCTCTGACTACTACTACCCAGAAATGTACCAGACCATCTGGACTG 3626
QY 1758 CTGGCAGAGACCCCAAGAGAGCCCAAGATTTGCAGACCTTTCGAAACACTAGTGA 1817
Db 3627 CTGGCATGAGACCCCAAGAGAGCCCTCGTTTTCAGAGTTGGTGGAGCATTTGGGAAA 3686
QY 1818 TTGCTTCAAGCAAAATGTAACAGAGATGGTAAAGACTACAT 1859
Db 3687 CCTCTGCAAGCAAAATGSCAGCAGAGTGGCAAGAGACTATAT 3728

Search completed: December 8, 2002, 09:38:18
Job time : 518.686 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2002, 07:52:47 ; Search time 2363.64 Seconds
(without alignments)
11586.276 Million cell updates/sec

Title: D64016_COPY_500_1440
Perfect score: 941
Sequence: 1 GCCGAAACACACGACAC.....CCCCGGTCITTCGCCGCGG 941

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_lm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vt.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_ru.*
20: em_or.*
21: em_ov.*
22: em_pat.*
23: em_ph.*
24: em_pi.*
25: em_ro.*
26: em_sts.*
27: em_un.*
28: em_vt.*
29: em_vt.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_higo_hum.*
40: em_higo_mus.*
41: em_higo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	941	100.0	1745	6	AR157385	AR157385 Sequence
2	941	100.0	1745	9	HOMES	D64016 Human gene
3	939.4	99.8	102119	9	ALI39005	ALI39005 Human DNA
4	667	70.9	840	6	E10168	E10168 Promoter of
5	613	65.1	6316	6	AX251105	AX251105 Sequence
6	613	65.1	6316	6	AX251868	AX251868 Sequence
7	613	65.1	6316	6	AX344256	AX344256 Sequence
8	613	65.1	6316	6	AX348647	AX348647 Sequence
9	613	65.1	6316	6	AX251106	AX251106 Sequence
10	510.8	54.3	6316	6	AX251869	AX251869 Sequence
11	510.8	54.3	6316	6	AX344257	AX344257 Sequence
12	510.8	54.3	6316	6	AX348648	AX348648 Sequence
13	510.8	54.3	6316	6	AX458548	AX458548 Sequence
14	510.8	54.3	6316	6	AX458549	AX458549 Sequence
15	375.4	39.9	174496	2	AC119074	AC119074 Rattus no
16	361.2	38.4	232575	2	AC122299	AC122299 Mus muscu
17	359.6	38.2	2648	10	MMU224863	MMU224863 Mus muscu
18	194.8	20.7	2313	6	BD003577	BD003577 Gene ther
19	194.8	20.7	2313	6	I82801	I82801 Sequence 5
20	194.8	20.7	2523	6	AR163540	AR163540 Sequence
21	194.8	20.7	2523	6	E13256	E13256 Human mRNA
22	194.8	20.7	2523	6	E14000	E14000 Human mRNA
23	194.8	20.7	2523	6	AR030842	AR030842 Sequence
24	194.8	20.7	2651	9	U01134	U01134 Human solub
25	194.8	20.7	7680	9	HSFLI	HSFLI Human flt m
26	130.6	13.9	3009	6	AX060540	AX060540 Sequence
27	78.2	8.3	159184	2	AC130633	AC130633 Rattus no
28	76.6	8.1	185263	2	AC129765	AC129765 Rattus no
29	76.6	8.1	174031	2	AC118303	AC118303 Rattus no
30	76.2	8.0	62847	2	AC125560	AC125560 Rattus no
31	75.6	7.9	95282	2	AC106630	AC106630 Rattus no
32	74.2	7.9	122113	2	AC126528	AC126528 Rattus no
33	74.2	7.8	292136	2	AC103544	AC103544 Rattus no
34	73	7.7	125159	2	AC127844	AC127844 Rattus no
35	72.4	7.7	174303	2	AC099104	AC099104 Rattus no
36	72	7.5	27913	2	AC129739	AC129739 Rattus no
37	71	7.5	111079	2	AC121738	AC121738 Rattus no
38	70.8	7.5	174031	2	AC118303	AC118303 Rattus no
39	70.8	7.5	72397	2	AC114443	AC114443 Rattus no
40	70.6	7.5	150238	2	CNS08CA9	AL831796 Oryza sat
41	70.6	7.5	155406	2	AC105514	AC105514 Rattus no
42	70.6	7.5	298166	2	AC087563	AC087563 Homo sapi
43	70.4	7.4	85022	2	AC121484	AC121484 Rattus no
44	70	7.4	123071	2	AC121481	AC121481 Rattus no
45	70	7.4	123071	2	AC121481	AC121481 Rattus no

ALIGNMENTS

RESULT 1
AR157385
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AR157385
Sequence
AR157385
AR157385.1
Unknown.
Unclassified.
1 (bases 1 to 1745)
Williams, L.I. and Morishita, K.
Promoter for VEGF receptor
Patent: US 6245512-A 1 12-JUN-2001;
Location/Qualifiers

1745 bp
DNA
linear
PAT 17-OCT-2001

Filed 3/15/99
date of filed
11/9/95

Pred. No. is the number of results predicted by chance to have a

source	1. 1745	315 a	501 c	621 g	308 t
BASE COUNT	315 a	501 c	621 g	308 t	
ORIGIN					
Query Match	100.0%	Score 941:	128 6:	Length 1745:	
Best Local Similarity	100.0%:	Pred. No. 1.2e-143:			
Matches 941:	Conservative	0:	Mismatches	0:	Indels
0:	Gaps	0:	Gaps	0:	Gaps
QY	1	GC	GA	AA	AA
DB	500	GC	GA	AA	AA
QY	61	GT	GC	CA	AA
DB	560	GT	GC	CA	AA
QY	121	CT	CG	GG	TG
DB	620	CT	CG	GG	TG
QY	181	GC	GG	GA	AG
DB	580	GC	GG	GA	AG
QY	241	AC	CG	TC	GC
DB	740	AC	CG	TC	GC
QY	301	GG	CG	AG	CT
DB	800	GG	CG	AG	CT
QY	361	GG	TC	GG	AG
DB	860	GG	TC	GG	AG
QY	421	AG	AG	GC	CG
DB	920	AG	AG	GC	CG
QY	481	AG	CA	CT	CC
DB	980	AG	CA	CT	CC
QY	541	CC	GT	AG	CC
DB	1040	CC	GT	AG	CC
QY	601	GG	TC	GG	AG
DB	1100	GG	TC	GG	AG
QY	661	AG	GT	TA	AA
DB	1160	AG	GT	TA	AA
QY	721	GC	GC	GC	GC
DB	1220	GC	GC	GC	GC
QY	781	CT	CG	GG	GT
DB	1280	CT	CG	GG	GT
QY	841	GT	TC	TC	TC
DB	1340	GT	TC	TC	TC
QY	901	AC	GA	AG	AG
DB	1400	AC	GA	AG	AG
QY	961	AG	GT	TA	AA
DB	1460	AG	GT	TA	AA

```

RESULT 2
HOMES      1745 bp   DNA      linear   PRI 14-APR-2000
LOCUS      Human gene for vascular endothelial growth factor receptor,
DEFINITION promoter and exon 1.
ACCESSION  D64016
VERSION    D64016.1 GI:1088437
KEYWORDS   vascular endothelial growth factor receptor.
SOURCE     Homo sapiens placenta DNA.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 1745)
AUTHORS    Morishita,K., Johnson,D.E. and Williams,L.T.
TITLE      A novel promoter for vascular endothelial growth factor receptor
           (flt-1) that confers endothelial-specific gene expression
JOURNAL    J. Biol. Chem. 270 (45), 27948-27953 (1995)
MEDLINE    96070934
REFERENCE  2 (bases 1 to 1745)
AUTHORS    Morishita,K.
TITLE      Direct Submission
JOURNAL    Submitted (31-AUG-1995) Kaoru Morishita, Daiichi Pharmaceutical
           Co., Ltd. Exploratory Research Laboratories 2; 1-16-13,
           Kita-Kasai, Edogawa-ku, Tokyo 134, Japan
           (Tel:03-3680-0151(ex.3722). Fax:03-5696-8718)
FEATURES   Location/Qualifiers
            1..1745
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="Placenta"
            1184..1171
            /evidence=not_experimental
            1196..1425
            /product="vascular endothelial growth factor receptor"
            /note="flt-1"
            /number=1
            /evidence=experimental
            1426..>1745
            /number=1
BASE COUNT 315 a 501 c 621 g 308 t
ORIGIN
Query: Match 100.0%; Score 941; DB 9; Length 1745;
      Local Similarity 100.0%; Pred. No. 1.2e-143;
      Matches 941; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Q: 1 GCGAAAAGACACGGACAGCTCCCTCGGACCTGACCTGGTTCGCAGTCTTCCCAAG 60
Db 500 GCGAAAAGACACGGACAGCTCCCTCGGACCTGACCTGGTTCGCAGTCTTCCCAAG 59
Q: 51 GTCCGACACAGCTCAGTTCCTCCCTCAGCGCTCCAGCTTCAGTGCCTGTGCGCAGGGT 120
Db 550 GTCCGACACAGCTCAGTTCCTCCCTCAGCGCTCCAGCTTCAGTGCCTGTGCGCAGGGT 619
Q: 121 CTCGGTGCTTCTTAGACTTCTCGGACAGTCTGAAGGGGTGAGAGCGCGGGACAGC 180
Db 620 CTCGGTGCTTCTTAGACTTCTCGGACAGTCTGAAGGGGTGAGAGCGCGGGACAGC 679
Q: 181 GCGGAGACAGCAGCGGAGACAGCGGACTCGCGCTCAGTCTCGTGCCTAAGAC 240
Db 680 GCGGAGACAGCAGCGGAGACAGCGGACTCGCGCTCAGTCTCGTGCCTAAGAC 739
Q: 241 ACCGTCCGCGAGGCGCGGACAGTTCCTTTGGATCGGACTTTCGCGCCCTAGGGCCAGGC 300
Db 740 ACCGTCCGCGAGGCGCGGACAGTTCCTTTGGATCGGACTTTCGCGCCCTAGGGCCAGGC 799
Q: 301 GCGGAGCTTCAGCCCTGTCCCTTCCCAAGTTCGCGGCGCCCGACAGCTCAGTAGCC 360
Db 800 GCGGAGCTTCAGCCCTGTCCCTTCCCAAGTTCGCGGCGCCCGACAGCTCAGTAGCC 859
Q: 351 GGGTGGAGGAGTCTGCAAGGATTTCTTGAGCGCATGGGACAGAGGCGGCAAGGSCA 420

```

Db 860 GGGTGGAGGAGCTCTCAAGGATTTCTGAGCCGATGGGAGGAGGAGGCGCAAGGCA 919
QY 421 AGAGGGCGCGGAGCAACACCTTGAACCTTCCGGGCGCGCTCCCGGGCCCGCTCGCC 480
Db 920 AGAGGGCGGAGCAACACCTTGAACCTTCCGGGCGCGCTCCCGGGCCCGCTCGCC 379
QY 481 AGACCTTCCCTACCGGCGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCT 540
Db 980 AGACCTTCCCTACCGGCGCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCT 1039
QY 541 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 1040 CCGTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
QY 601 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 560
Db 1100 GGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1159
QY 661 AAGGTTAAATGCGCGCGCGCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Db 1160 AAGGTTAAATGCGCGCGCGCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1219
QY 721 GCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 1220 GCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1279
QY 781 CTCGGGGCTCGGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 840
Db 1280 CTCGGGGCTCGGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 1339
QY 841 GGTGTCT 900
Db 1340 GGTGTCT 1399
QY 901 ACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 941
Db 1400 ACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

RESULT 3
AL139005/c
LOCUS Human DNA sequence from clone RP11-95G6 on chromosome 13, complete
DEFINITION sequence.
ACCESSION AL139005
VERSION AL139005.12 GI:14018246
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 102119)
Direct Submission
Submitted (09-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clones@sanger.ac.uk
On May 11, 2001 this sequence version replaced gi:13274245.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: EMBL, EMBL; SW,

SWISSPROT: Tr, TREMBL, Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr13
RP11-95G6 is from the library RP11-11.1 constructed by the group of
Pierle de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.5
IMPORTANT: This sequence is not the entire insert of clone
RP11-95G6. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-95G6 is at 102119 in this
sequence. The true left end of clone RP11-57H24 is at 59065 in this
sequence. The true right end of clone RP11-502P18 is at 100 in this
sequence.

FEATURES
source
1..102119
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-95G6"
/clone_lib="RP11-11.1"
851..979
repeat_region
/note="FLAM_A repeat: matches 1..129 of consensus"
repeat_region
1038..1087
/note="25 copies 2 mer gt 92% conserved"
repeat_region
1044..1087
/note="11 copies 4 mer ggtg 100% conserved"
repeat_region
2014..2102
/note="12 repeat: matches 2645..2742 of consensus"
repeat_region
3151..3398
/note="MIR repeat: matches 2..252 of consensus"
repeat_region
3745..3813
/note="3 copies 23 mer 82% conserved"
repeat_region
4238..4480
/note="AluSq repeat: matches 1..243 of consensus"
repeat_region
4481..4512
/note="8 copies 4 mer aaat 93% conserved"
repeat_region
5434..5467
/note="17 copies 2 mer aa 82% conserved"
repeat_region
6041..6085
/note="9 copies 5 mer ttatg 84% conserved"
repeat_region
6384..6387
/note="AluSx repeat: matches 1..300 of consensus"
repeat_region
7056..7367
/note="AluSg repeat: matches 1..310 of consensus"
repeat_region
8032..8343
/note="AluSx repeat: matches 1..312 of consensus"
repeat_region
8585..8771
/note="MER5B repeat: matches 84..178 of consensus"
repeat_region
9322..9957
/note="12 copies 3 mer tta 94% conserved"
repeat_region
9358..10079
/note="AluJo/FRAM repeat: matches 156..282 of consensus"
repeat_region
10539..10707
/note="L1M25 repeat: matches 7772..7958 of consensus"
repeat_region
10739..10916
/note="L1M24 repeat: matches 7645..7820 of consensus"
repeat_region
12636..12928
/note="Aluub repeat: matches 1..307 of consensus"
repeat_region
13961..14177
/note="MIR repeat: matches 46..252 of consensus"
repeat_region
14564..14862
/note="AluSp repeat: matches 1..305 of consensus"
repeat_region
15571..15708
/note="6 copies 23 mer 70% conserved"
repeat_region
15756..15701
/note="63 copies 2 mer tt 71% conserved"
repeat_region
15890..15898
/note="7 copies 17 mer 74% conserved"

```
repeat_region 15625..15704
/note="20 copies 4 mer tctt 87% conserved"
repeat_region 15708..16014
/note="AluX repeat: matches 1..308 of consensus"
repeat_region 16022..16126
/note="AluSg/x repeat: matches 202..306 of consensus"
repeat_region 17111..17199
/note="MIR repeat: matches 1..94 of consensus"
repeat_region 17832..17897
/note="MIR repeat: matches 193..260 of consensus"
repeat_region 17845..18023
/note="L2 repeat: matches 2567..2739 of consensus"
repeat_region 18059..18137
/note="3 copies 23 mer 87% conserved"
repeat_region 18494..18777
/note="AluSp repeat: matches 3..285 of consensus"
repeat_region 19334..19385
/note="L2 repeat: matches 2648..2701 of consensus"
repeat_region 19350..19582
/note="MIR repeat: matches 2..241 of consensus"
repeat_region 19810..19971
/note="MIR repeat: matches 85..258 of consensus"
repeat_region 20842..20921
/note="MIR repeat: matches 54..140 of consensus"
repeat_region 20366..21034
/note="FLAM_A repeat: matches 51..128 of consensus"
repeat_region 21009..21072
/note="AluA8 repeat: matches 134..210 of consensus"
repeat_region 21868..22156
/note="AluB repeat: matches 1..302 of consensus"
repeat_region 22892..23190
/note="AluY repeat: matches 1..299 of consensus"
repeat_region 25148..25236
/note="MIR repeat: matches 52..147 of consensus"
repeat_region 27232..27532
/note="AluSg repeat: matches 1..313 of consensus"
repeat_region 29375..29442
/note="34 copies 2 mer tt 66% conserved"
repeat_region 29702..30027
/note="L2 repeat: matches 2356..2710 of consensus"
repeat_region 31857..31999
/note="MIR repeat: matches 25..141 of consensus"
repeat_region 35931..36016
/note="MER3 repeat: matches 1..208 of consensus"
repeat_region 37895..38193
/note="AluSg repeat: matches 1..299 of consensus"
repeat_region 40064..40284
/note="MIR repeat: matches 14..241 of consensus"
repeat_region 40693..40743
/note="L2 repeat: matches 2647..2699 of consensus"
repeat_region 42794..43196
/note="MIR repeat: matches 3..420 of consensus"
repeat_region 43395..43657
/note="MIR repeat: matches 77..140 of consensus"
repeat_region 43937..44142
/note="MIR repeat: matches 14..262 of consensus"
repeat_region 44173..44831
/note="MER50 repeat: matches 1..709 of consensus"
repeat_region 44354..45088
/note="FLAM_C repeat: matches 2..129 of consensus"
repeat_region 45992..46035
/note="22 copies 2 mer ac 90% conserved"
repeat_region 45996..46035
/note="10 copies 4 mer acac 95% conserved"
repeat_region 47173..47453
/note="L1MB8 repeat: matches 2311..2594 of consensus"
repeat_region 47454..47743
/note="AluX repeat: matches 1..292 of consensus"
repeat_region 47744..48488
/note="L1MB8 repeat: matches 2594..3347 of consensus"
repeat_region 48489..48761
/note="AluSx repeat: matches 1..275 of consensus"
repeat_region 48762..50491
```

```
repeat_region 50492..50780
/note="L1MB8 repeat: matches 3347..5028 of consensus"
repeat_region 50781..51024
/note="AluB repeat: matches 1..285 of consensus"
repeat_region 51044..51179
/note="8 copies 17 mer 66% conserved"
repeat_region 51190..52061
/note="L1MB8 repeat: matches 5271..6171 of consensus"
repeat_region 52293..52577
/note="AluJ repeat: matches 1..295 of consensus"
repeat_region 53786..53900
/note="5 copies 23 mer 68% conserved"
repeat_region 53786..53875
/note="22 copies 4 mer agga 68% conserved"
repeat_region 54044..54138
/note="MIR repeat: matches 46..142 of consensus"
repeat_region 54435..54742
/note="AluSx repeat: matches 1..308 of consensus"
repeat_region 54900..54983
/note="42 copies 2 mer gg 64% conserved"
repeat_region 55026..55061
Query Match 99.8%; Score 939.4; DB 9; Length 1021119;
Best Local Similarity 99.9%; Pred. No. 7.8e-144;
Matches 940; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGACGTCTTCCCAAG 60
Db 99118 GCCGAAAAGACAGGACAGCTCCCTGGGACCTGAGCTGGTTCGACGTCTTCCCAAG 99059

Qy 51 GTCCCAACAGCGTCAGTCCCTCAGGCTCCAGTTCAGTTCAGTTCGCGAGGT 120
Db 99058 GTCCCAACAGCGTCAGTTCCTCAGGCTCCAGTTCAGTTCAGTTCGCGAGGT 98999

Qy 121 CTCGGTTCCTTCAGACTTCCTGGGACAGTCTGAAGGGTCAGGAGCGGCGACAGC 180
Db 98998 CTCGGTTCCTTCAGACTTCCTGGGACAGTCTGAAGGGTCAGGAGCGGCGACAGC 98939

Qy 181 GCGGAAGACAGGACAGGAGGACAGCGGCTCCGCTCAGTTCCTCGTTCGCAAGAC 240
Db 98938 GCGGAAGACAGGACAGGAGGACAGCGGCTCCGCTCAGTTCCTCGTTCGCAAGAC 98879

Qy 241 ACCGTTCGGGAGCGCGGCTTCCTTCGATCGGACTTCGCGCTTCAGGCGCAGGC 300
Db 98878 ACCGTTCGGGAGCGCGGCTTCCTTCGATCGGACTTCGCGCTTCAGGCGCAGGC 98819

Qy 301 GCGGAGCTTCAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 360
Db 98818 GCGGAGCTTCAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 98759

Qy 361 GCGTGGAGGAGTTCGAAGGATTCCTCAGCGGATTCGCGAGGAGGAGGCAAGGCA 420
Db 98758 GCGTGGAGGAGTTCGAAGGATTCCTCAGCGGATTCGCGAGGAGGAGGCAAGGCA 98599

Qy 421 AGAGGCGCGGAGCAAGACCTTCGCGGCGCGGCTTCGCGGCGCGCTTCGCGCGC 480
Db 98598 AGAGGCGCGGAGCAAGACCTTCGCGGCGCGGCTTCGCGGCGCGCTTCGCGCGC 98639

Qy 481 AGACCTTCCTCAGCGGCTTCGCGCGCGGCGCGGCTTCGCGCGCGCTTCGCGCGC 540
Db 98638 AGACCTTCCTCAGCGGCTTCGCGCGCGGCGCGGCTTCGCGCGCGCTTCGCGCGC 98579

Qy 541 CCGTAGCGCGGAGGAGCGCTTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 98578 CCGTAGCGCGGAGGAGGAGCGCTTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 98519

Qy 501 GGGTGGGAGCGCTTCGAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 98518 GGGTGGGAGCGCTTCGAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 98459

Qy 561 AGGTTATTAATCGCGCGCGCTTCGCTTCATCGAGGTCGCGGAGGAGGAGGAGGAGG 720
Db 98459 AGGTTATTAATCGCGCGCGCTTCGCTTCATCGAGGTCGCGGAGGAGGAGGAGGAGG 98399
```

Db	98458	AAGGTATAAAATGCCCCCGCCCTCGGCTGCTCTTCAAGAGTCCGCGGAGAGCTCGA	98399
QY	721	GGCGCCAGCGGACACTCTCTGGCTCTCTCCGSCAGCGCGCGCTCGAGCGGG	780
Db	98398	GGCGCCAGCGGACACTCTCTGGCTCTCTCCGSCAGCGCGCGCTCGAGCGGG	98339
QY	781	CTCCGGGCTCGGTCAGCGCGCCAGCGGCGCTCGCGCGAGGATACCGGCGGAGT	840
Db	98338	CTCCGGGCTCGGTCAGCGCGCCAGCGGCGCTCGCGCGAGGATACCGGCGGAGT	98279
QY	841	GGTGTCTCTGGCTCGAGCGCGAGCGGCGCTCAGCGCGCGCGCGCGCGGCGA	900
Db	98278	GGTGTCTCTGGCTCGAGCGCGAGCGGCGCTCAGCGCGCGCGCGCGCGGCGGA	98219
QY	901	ACGAGGACGAGTCTCTCGCGCGCGGCTTTTGGCGCGCGG 941	
Db	98218	ACGAGGACGAGTCTCTCGCGCGCGGCTTTTGGCGCGCGG 98178	
SULT 4			
3168			
CUS			
DEFINITION	E10168	840 bp DNA linear	PAT 23-SEP-1997
ACCESSION	E10168	Promoter of human flt gene encoding receptor type tyrosine kinase.	
VERSION	E10168.1	GI:22025997	
KEYWORDS	JP 1995289263-A/1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Shibuya.M.		
JOURNAL	PROMOTER FOR GENE MANIFESTATION		
COMMENT	Patent: JP 1995289263-A 1 07-NOV-1995;		
	DAI ICHI SEIYAKU CO LTD		
	OS Homo sapiens (human)		
	PN JP 1995289263-A/1		
	PD 07-NOV-1995		
	PF 22-APR-1994		
	PI SHIBUYA MASASHI		
	PC C12N15/09,A61K31/70,A61K45/00,C12N5/10,C12P21/02//C12P21/04,		
	C12N1/68		
	PC (C12N5/10,C12P1:91),(C12P21/02,C12P1:91);		
	CC strandedness: Double;		
	CC topology: Linear;		
	CC hypothetical: No;		
	CC anti-sense: No;		
PH	Key	Location/Qualifiers	
FI	Source	1..840	
FI	/organism='Homo sapiens'		
FI	/tissue_type='placenta'		
FI	promoter	1<..489	
FI	TATA_signal	457..462	
FI	CDS	771..840	
FI	/product='receptor type tyrosine kinase'		
FI	misc_feature	343..347	
FI	/note='Ets core motif'		
FI	misc_feature	350..354	
FI	/note='Ets core motif'		
FI	misc_feature	435..440	
FI	/note='Ets core motif'		
FI	Location/Qualifiers		
source	1..840		
BASE COUNT	120 a	269 c	336 g
ORIGIN	115 t		
Query Match	70.9%	Score 567;	DB 5; Length 840;
Best Local Similarity	98.6%	Pred. No. 4.9e-99;	
Matches 725; Conservative	0;	Mismatches 5;	Indels 5; Gaps 5;

QY	208	CCGAGCTCGCGCTCAGTCTCGTCCGAGACACCGTCCGAGGCGCGCGCGCTTCC	267
Db	1	CCGAGCTCGCGCTCAGTCTCGTCCGAGACACCGTCCGAGGCGCGCGCGCTTCC	60
QY	268	CTTGAATCGGACATTCGCGCCCTAGGCGCAGCGCGGAGCTTCAACCTTGTCCCTCC	327
Db	61	CTTGAATCGGACATTCGCGCCCTAGGCGCAGCGCGGAGCTTCAACCTTGTCCCTCC	118
QY	328	CAGTTTCGGCGCGCGCCCGACAGCTGAGTAAAGCGCGGTGGAGGAGTCTCCAGGATTC	387
Db	119	CAGTTTCGGCGCGCGCCCGACAGCTGAGTAAAGCGCGGTGGAGGAGTCTCCAGGATTC	178
QY	388	TCAGCGCATGGCGCAGCAGGCGCAGGCGCAGGCGCAGGCGCGGAGCAAGACCTGAAC	447
Db	179	TCAGCGCATGGCGCAGGCGCAGGCGCAGGCGCAGGCGCGGAGCAAGACCTGAAC	237
QY	449	CTCCCGGCGCGCGCTCCCGCGCGCGCTGCGCAGCAGCTCCCGCAGCGCGCTCGGCCCC	507
Db	238	CTCCCGGCGCGCGCTCCCGCGCGCGCTGCGCAGCAGCTCCCGCAGCGCGCTCGGCCCC	297
QY	508	GGCGCATCGCGCTCGTCCGCGCGCGCGCTCGTCCGAGCGCGGAGGCGAGCGCTGG	566
Db	298	GGCGCATCGCGCTCGTCCGCGCGCGCGCTCGTCCGAGCGCGGAGGCGAGCGCTGG	357
QY	567	GAGCAAGAGAGGAGTAGTGGGAGCGGATGAGGGGTGGGGACCCCTTGAGTCAACA	626
Db	358	GAGCAAGAGAGGAGTAGTGGGAGCGGATGAGGGGTGGGGACCCCTTGAGTCAACA	417
QY	627	GAGGAGGTCGCGGGTAGGAGTGGGCTGGGGAAGGTATTAATCGCCCGCGCTCG	686
Db	418	GAGGAGGTCGCGGGTAGGAGTGGGCTGGGGAAGGTATTAATCGCCCGCGCTCG	477
QY	687	GCTGCTCTTCATCGAGGTCCCGGGAGGCTCGGAGCGCGGCTCGGCGGACACTCTCGG	746
Db	478	GCTGCTCTTCATCGAGGTCCCGGGAGGCTCGGAGCGCGGCTCGGCGGACACTCTCGG	537
QY	747	CTCTCTCTCGCGCGCGCGCTCGGAGCGGCTCGGAGCGGCTCGGCGGCTCGGCGGCG	806
Db	538	CTCTCTCTCGCGCGCGCGCTCGGAGCGGCTCGGAGCGGCTCGGCGGCTCGGCGGCG	597
QY	807	CGGCGCTCGGCGCGGAGTACCGGGGAGGTGGTGTCTCTCTGCTCGGCGCGCG	866
Db	598	CGAGCG-CTGGCGGCGGAGTATCCCGGGGAGTGGTGTCTCTCTGCTCGGCGCGCG	656
QY	867	ACGCGGCTCAGGCG	926
Db	557	ACGCGGCTCAGGCG	926
QY	927	GCTTTGGCGCGCGG 941	
Db	717	GTCGTTGGCGCGGG 731	
RESULT 5			
AX251105	AX251105	6316 bp DNA linear	PAT 05-OCT-2001
LOCUS	Sequence 73 from Patent WO0168912.		
DEFINITION	AX251105		
ACCESSION	AX251105		
VERSION	AX251105.1	GI:15984528	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	artificial construct		
REFERENCE	1 (bases 1 to 6316)		
AUTHORS	Olek A., Piepenbrock, C. and Berlin, K.		
TITLE	Diagnosis of diseases associated with tumor suppressor genes and		
JOURNAL	Patent: WO 0168912-A 73 20-SEP-2001;		
FEATURES	oncodenes		
Location/Qualifiers	Epigenomics AG (DE)		
1..6316			
/organism='synthetic construct'			
/db_xref='taxon:32630'			

3

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

DB 30630 TCITTCGCTATGCCAGA 30614

RESULT 3

AC128409/c

LOCUS

DEFINITION

AC128409

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 144090)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaralungu,H.C., Are,J.R., Ayele,M., Banks,J.,

Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonica,D.,

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.I., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.B.,

Douthwater,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Litcharge,O., Lieu,C., Liu,J., Liu,M., Louissegh,H.,

Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,S.,

Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,K., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,

Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,

Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,

Williams,G., Williamson,A., Wiczek,R., Wooden,S., Worley,K.,

Wu,C., Wu,T., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 144090)

Worley,K.C.

Direct Submission

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYFL

Center clone name: CH230-448N2

Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.950325

Consensus quality: 101711 bases at least Q40

Consensus quality: 106747 bases at least Q30

Consensus quality: 110016 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 57 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1636: contig of 1636 bp in length

* 1637 1736: gap of unknown length

* 1737 3403: contig of 1667 bp in length

* 3404 3503: gap of unknown length

* 3504 4554: contig of 1051 bp in length

* 4555 4654: gap of unknown length

* 4655 5864: contig of 1210 bp in length

* 5865 5964: gap of unknown length

* 5965 7115: contig of 1151 bp in length

* 7116 7215: gap of unknown length

* 7216 8510: contig of 1295 bp in length

* 8511 8610: gap of unknown length

* 8611 9955: contig of 1345 bp in length

* 9956 10055: gap of unknown length

* 10056 11380: contig of 1325 bp in length

* 11381 11480: gap of unknown length

* 11481 12607: contig of 1127 bp in length

* 12608 12707: gap of unknown length

* 12708 13653: contig of 1156 bp in length

* 13654 13963: gap of unknown length

* 13964 13963: gap of unknown length

* 13964 13963: gap of unknown length

* 13963 14621: gap of unknown length

* 14622 17024: contig of 1562 bp in length

* 17025 17124: gap of unknown length

* 17125 18879: contig of 1755 bp in length

* 18880 18979: gap of unknown length

* 18980 20913: contig of 1934 bp in length

* 20914 21013: gap of unknown length

* 21014 2216: contig of 2203 bp in length

* 2217 2316: gap of unknown length

* 2317 23554: contig of 2238 bp in length

* 23555 25654: gap of unknown length

* 25655 27062: contig of 1408 bp in length

* 27063 27162: gap of unknown length

* 27163 28572: contig of 1410 bp in length

* 28573 28672: gap of unknown length

* 28673 29930: contig of 1258 bp in length

* 29931 30030: gap of unknown length

* 30031 32527: contig of 2457 bp in length

* 32528 32627: gap of unknown length

* 32628 33881: contig of 1254 bp in length

* 33882 33981: gap of unknown length

* 33982 36613: contig of 2632 bp in length

* 36614 36713: gap of unknown length

* 36714 38530: contig of 1817 bp in length

* 38531 38630: gap of unknown length

* 38631 40221: contig of 1591 bp in length

* 40222 40321: gap of unknown length

* 40322 42216: contig of 1895 bp in length

* 42217 42316: gap of unknown length

* 42317 44902: contig of 2586 bp in length

* 44903 45002: gap of unknown length

* 45003 47520: contig of 2518 bp in length

* 47521 47620: gap of unknown length

* 47621 49384: contig of 1764 bp in length


```

gene      1. .4017
          /gene="FLT1"
CDS       1. .4017
          /gene="FLT1"
          /note="VEGFR; tyrosine kinase"
          /codon_start=1
          /product="vascular endothelial growth factor receptor"
          /protein_id="AAC16449.1"
          /db_xref="GI:3132831"
          /translation="MVSVDGTGVLICALLSLLIGSSSSGSKLKPPELSIKGTOHMQ
          AGTLLHOCGEAAHKSLEPMVSKESSELSITKACGRNGKQFCSILINIAQAHNI
          GYSCKYLAVPTSKKTESAIYIFISDGRFFVEMSEIPEIHTMTEGRELIVPCRV
          TSPNITVAKFPDITLIPDKRIIWDKRGFIISNATYKEIGLLTCEATVNGHYKRT
          NLTHTQNTIIVQIISTPPVKLLRGHTLINCTATPLNTRVQMTSPDEKMKRA
          SVRRIDNSHANIFYSLTIDKQNDKGLYICRVSRGSPKSVNTSVHYDKAFI
          TVKHKQVLTAVGKSYLSMKVKAPEVWVWLDGLPATEKSAVLYTRGSLII
          KQVTEAGNYITLLSKQSNVFNKLTATLVNWKPOIYKAVSPDPALYLGSSQ
          ILTCYAGIPOPTIKWPCNHNHSEAKDCFCNNEESFILDASNMGNRJESIQ
          PTEGEDKLSCTVNFYRDVWILLIRVNNRIMHYSISKMAITKEHSITLNIIM
          NVSLQDSGYACRARNVYIGELLOKKEITIDQEAFLNLSHIVALSSEITLQ
          HANGPEQITWFKNNHIOQEPGIIIGPGSTFIERVTEDEGVYHCKAIKQGSV
          ESSAYLTVOGSDSKSNLELITCTQVAAITLWILLIRKMRSSSEIKIDYLSII
          MDPDEVPDQCEBPLVDASKNEFARERKLGKSLRGAGFKVQVQASAFGIKASCTIC
          TVAVMLKEGATSEYKALMTELKILTHIGHLVNALLGACTKGGPLWIVYECYK
          GNLNYSKSRDLEFFLNKDAALHMEPKKEPGLQEGKPRDVSSESPASSGFO
          EDKLSVDEEEDSDGYKEPITMEDLISYFQVARGMEFLSRKCIHRLAARNILL
          SENNVKICDFGLARDIYNPDYVRKGTIRPLKAWAPESIFDKIYISKSDVNSYGLV
          LNEIFSLGSPYPCVOMDEDFCSLRGEMRAREYSTPIYQIMLDCHWRPKERPR
          FAELVEKGLDLOANVQDGKQYIPINAILTGNSSFTYSPAFSEDFKESISAPFN
          SCSSDDVYVNAFKENSLERIKTFEELLPNATSMFDYQDSDSTILLASPLMKRTWTD
          SAPKASLKIDLVTSKESGSLDSVRSFCHSSCGHVSEGRKRTFYDHAELERKIAK
          CSPPPDYNVWLVSTPPI"
BASE COUNT 1236 a 915 c 903 g 563 t
ORIGIN

Query Match      38.0%; Score 114.4; DB 9; Length 4017;
Best Local Similarity 99.1%; Pred. No. 3.5e-23;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 66 AGACTCGACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 125
Db 3173 ATATCGACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 5232
QY 126 CCAAGACGACGTGTGGTCTTACGAGATGATGCTGTGGGAAATCTTCCCTAGGT 181
Db 3233 CCAAGACGACGTGTGGTCTTACGAGATGATGCTGTGGGAAATCTTCCCTAGGT 3288

RESULT 6
HSLT
LOCUS      7680 bp mRNA linear PRI 15-NOV-1993
DEFINITION Human flt mRNA for receptor-related tyrosine kinase.
ACCESSION X51602
VERSION   X51602.1 GI:31431
KEYWORDS  flt gene; fms-related tyrosine kinase gene; tyrosine kinase.
SOURCE    Homo sapiens.
ORGANISM  Homo sapiens
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 7680)
           Shibuya M.
           Direct Submission
           Submitted (02-JAN-1989) Shibuya M., Institute of Medical Science,
           University of Tokyo, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108,
           Japan
           2 (bases 1 to 7680)
           Shibuya M., Yamaguchi S., Yamane A., Ikeda T., Tojo A.,
           Matsushima H. and Sato M.
           Nucleotide sequence and expression of a novel human receptor-type
           tyrosine kinase gene (flt) closely related to the fms family
           Oncogene 5 (4), 519-524 (1990)
           90221591

```

```

PUBMED    2158033
REFERENCE 3 (bases 1 to 7680)
AUTHORS   Han, H.-J., Fujiwara, I., Shin, S. and Nakamura, Y.
TITLE     Dinucleotide repeat polymorphism in the 3' non-coding region of the
          FII gene
JOURNAL   Hum. Mol. Genet. 2 (12), 2204 (1993)
MEDLINE   94154724
PUBMED    8111405
COMMENT   Data kindly reviewed (20-JUL-1990) by Shiduya M.
FEATURES   Location/Qualifiers
           1..7680
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /chromosome="13"
           /map="13q12"
           /clone="3-7, 3-5"
           /tissue_type="placenta"
           250..4266
           /note="flt gene product (AA 1-1336)"
           /codon_start=1
           /protein_id="CAA35946.1"
           /db_xref="GI:31432"
           /db_xref="SWISS-PROT:P17948"
           /translation="MVSVDGTGVLICALLSLLIGSSSSGSKLKPPELSIKGTOHMQ
           AGTLLHOCGEAAHKSLEPMVSKESSELSITKACGRNGKQFCSILINIAQAHNI
           GYSCKYLAVPTSKKTESAIYIFISDGRFFVEMSEIPEIHTMTEGRELIVPCRV
           TSPNITVAKFPDITLIPDKRIIWDKRGFIISNATYKEIGLLTCEATVNGHYKRT
           NLTHTQNTIIVQIISTPPVKLLRGHTLINCTATPLNTRVQMTSPDEKMKRA
           SVRRIDNSHANIFYSLTIDKQNDKGLYICRVSRGSPKSVNTSVHYDKAFI
           TVKHKQVLTAVGKSYLSMKVKAPEVWVWLDGLPATEKSAVLYTRGSLII
           KQVTEAGNYITLLSKQSNVFNKLTATLVNWKPOIYKAVSPDPALYLGSSQ
           ILTCYAGIPOPTIKWPCNHNHSEAKDCFCNNEESFILDASNMGNRJESIQ
           MAIECKKMASTLVADSRISGIYICIASKNVGTGNISFYITDVPNGFHNLEK
           PTEGEDKLSCTVNFYRDVWILLIRVNNRIMHYSISKMAITKEHSITLNIIM
           NVSLQDSGYACRARNVYIGELLOKKEITIDQEAFLNLSHIVALSSEITLQ
           HANGPEQITWFKNNHIOQEPGIIIGPGSTFIERVTEDEGVYHCKAIKQGSV
           ESSAYLTVOGSDSKSNLELITCTQVAAITLWILLIRKMRSSSEIKIDYLSII
           MDPDEVPDQCEBPLVDASKNEFARERKLGKSLRGAGFKVQVQASAFGIKASCTIC
           TVAVMLKEGATSEYKALMTELKILTHIGHLVNALLGACTKGGPLWIVYECYK
           GNLNYSKSRDLEFFLNKDAALHMEPKKEPGLQEGKPRDVSSESPASSGFO
           EDKLSVDEEEDSDGYKEPITMEDLISYFQVARGMEFLSRKCIHRLAARNILL
           SENNVKICDFGLARDIYNPDYVRKGTIRPLKAWAPESIFDKIYISKSDVNSYGLV
           LNEIFSLGSPYPCVOMDEDFCSLRGEMRAREYSTPIYQIMLDCHWRPKERPR
           FAELVEKGLDLOANVQDGKQYIPINAILTGNSSFTYSPAFSEDFKESISAPFN
           SCSSDDVYVNAFKENSLERIKTFEELLPNATSMFDYQDSDSTILLASPLMKRTWTD
           SAPKASLKIDLVTSKESGSLDSVRSFCHSSCGHVSEGRKRTFYDHAELERKIAK
           CSPPPDYNVWLVSTPPI"
BASE COUNT 2279 a 1661 c 1739 g 2001 t
ORIGIN

Query Match      37.5%; Score 112.8; DB 9; Length 7680;
Best Local Similarity 98.3%; Pred. No. 1e-22;
Matches 114; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 56 AGACTCGACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 125
Db 3422 ATATCGACTTCTCTGAAATGATGCTCTGAAATCTATCTTIGACAAAATCTACAGCA 3481
QY 126 CCAAGACGACGTGTGGTCTTACGAGATGATGCTGTGGGAAATCTTCCCTAGGT 181
Db 3482 CCAAGACGACGTGTGGTCTTACGAGATGATGCTGTGGGAAATCTTCCCTAGGT 3537

RESULT 7
HSLT
LOCUS      HSA312688
DEFINITION Homo sapiens chromosome 3 clone RP6-329g3 map 3p21.3, ***
ACCESSION AJ312688
VERSION   AJ312688.2 GI:13559235
KEYWORDS  HIG; HIGS_PHASE2.
SOURCE    human.
ORGANISM  Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1
Kiss, H., Yang, Y., Kiss, C., Andersson, K., Klein, G., Imren, S. and
Dumanski, J. P.
The transcriptional map of the common eliminated region 1 (CICER1)
in 3p21.3
Eur. J. Hum. Genet. 10 (1), 52-61 (2002)
15906202
PUBMED
11896456
REFERENCE
2 (bases 1 to 220965)
Kiss, H.
Direct Submission
Submitted (01-APR-2001) Kiss H., Microbiology and Immunobiology
Center (MTC), Karolinska Institute, Box 280, Stockholm, S-17177,
SWEDEN
On Apr 5, 2001 this sequence version replaced gi:1354653.
The sequence is a consensus sequence of clone RP6-787C23 (1-140400
bp),
clone RP6-32923 (31212-220965 bp), clone RP6-14661 (partially,
1-6800 bp)
and clone RP6-18911 (partially, 1-108303 bp). The sequencing
contigs are
in order and the gaps between them are represented by 100 Ns.
Contig 1:
1-11731 bp Contig 2: 11832-26218 bp Contig 3: 26319-28347 bp
Contig 4:
28448-42160 bp Contig 5: 42261-55059 bp Contig 6: 55160-61578 bp
Contig 7: 61679-97342 bp Contig 8: 97443-117655 bp Contig 9:
117756-118727 bp
Contig 10: 118828-121834 bp Contig 11: 121935-127855 bp Contig
12:
127956-129383 bp Contig 13: 129484-131747 bp Contig 14:
131848-132316 bp
Contig 15: 132417-134455 bp Contig 16: 134556-135527 bp Contig
17:
135628-189051 bp Contig 18: 189152-189476 bp Contig 19:
189577-191375 bp
Contig 20: 191476-201473 bp Contig 21: 201574-202307 bp Contig
22:
202408-204878 bp Contig 23: 204979-213531 bp Contig 24:
213632-218109 bp
Contig 25: 218210-219800 bp Contig 26: 219901-220965 bp.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 11731: contig of 11731 bp in length
* 11732 11831: gap of 100 bp
* 11832 26218: contig of 14387 bp in length
* 26219 26318: gap of 100 bp
* 26319 28347: contig of 2029 bp in length
* 28348 28447: gap of 100 bp
* 28448 42160: contig of 13713 bp in length
* 42161 42360: gap of 200 bp
* 42361 55059: contig of 12899 bp in length
* 55060 55159: gap of 100 bp
* 55160 61578: contig of 6419 bp in length
* 61579 61678: gap of 100 bp
* 61679 97342: contig of 35664 bp in length
* 97343 97442: gap of 100 bp
* 97443 117655: contig of 20213 bp in length
* 117656 117755: gap of 100 bp
* 117756 118727: contig of 972 bp in length
* 118728 118827: gap of 100 bp
* 118828 121834: contig of 3007 bp in length
* 121835 121934: gap of 100 bp

* 121935 127855: contig of 5921 bp in length
* 127856 127955: gap of 100 bp
* 127956 129383: contig of 1428 bp in length
* 129384 129483: gap of 100 bp
* 129484 131747: contig of 2264 bp in length
* 131748 131847: gap of 100 bp
* 131848 132316: contig of 489 bp in length
* 132317 132416: gap of 100 bp
* 132417 134455: contig of 2039 bp in length
* 134456 134555: gap of 100 bp
* 134556 135527: contig of 972 bp in length
* 135528 135627: gap of 100 bp
* 135628 189051: contig of 53424 bp in length
* 189052 189151: gap of 100 bp
* 189152 189476: contig of 325 bp in length
* 189477 189576: gap of 100 bp
* 189577 191375: contig of 1799 bp in length
* 191376 191475: gap of 100 bp
* 191476 201473: contig of 9998 bp in length
* 201474 201573: gap of 100 bp
* 201574 202307: contig of 734 bp in length
* 202308 204878: gap of 100 bp
* 204879 204978: gap of 100 bp
* 204979 213531: contig of 8553 bp in length
* 213532 213631: gap of 100 bp
* 213632 218109: contig of 4478 bp in length
* 218110 218209: gap of 100 bp
* 218210 219800: contig of 1591 bp in length
* 219801 219900: gap of 100 bp
* 219901 220965: contig of 1065 bp in length.

FEATURES
Location/Qualifiers
1..220965
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p21.3"
/clone="RP6-32923"
BASE COUNT 62577 a 41674 c 45661 g 65453 t 2600 others
ORIGIN

Query Match 34.7%; Score 104.4; DB 2; Length 220965;
Best Local Similarity 94.7%; Pred. No. 3.4e-20;
Matches 108; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 68 ACACGACCTCTCTGAAATGGATGGCTCTGCAATCTATCTTTGACAAAATCTACAGCACC 127
|||||
DB 24786 ACTTGCCTCTCTGAAATGGATGGCTCTTGAATCTATCTTTGACAAAATCTACAGCACC 24727
|||||

QY 128 AAGAGGACGCTGTGTCTAGGAGTATCTGTGGGAAATCTCTCTTASGT 181
|||||

DB 24726 AAGAGGACGCTGTGTCTAGGAGTATCTGTGGGAAATCTCTCTTASGT 24673
|||||

RESULTS
AC104435/c
LOCUS AC104439 197275 bp DNA linear PRI 20-JUN-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-793E15, complete sequence.
ACCESSION AC104439 AC024739
VERSION AC104439.2 GI:21490240
KEYWORDS H.G.
SOURCE H.G.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 197275)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Sankhramachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
et al; Haugen, E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197279)
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.

TITLE
JOURNAL

Direct Submission
Submitted (11-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 197279)

REFERENCE
AUTHORS

Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.

TITLE
JOURNAL

Direct Submission
Submitted (20-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jun 20, 2002 this sequence version replaced gi:1748621.

COMMENT

Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgts@u.washington.edu
Drafting Center: WUGSC

Project Information
Center project name: chr-3
Center clone name: RP11-793E15 (bc0564)

Summary Statistics
Sequencing vector: unknown; 52% of reads
Sequencing vector: plasmid; 108752; 48% of reads
Chemistry: Dye-terminator ET; 94% of reads
Chemistry: Dye-terminator Big Dye; 6% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 197168 bases at least Q40
Consensus quality: 197255 bases at least Q30
Consensus quality: 197275 bases at least Q20
Insert size: 197279; sum-of-contigs
Quality coverage: 8.2x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP11-91E8 (UWGC:bc0216) AC026349
3': CTD-2563A18 (UWGC:bc0730)

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

HindIII ----- EcoRI

SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt
2687	2617	8949	8586	8696	8661

6382	6410	2067	2160	6	<860
512	<800	7846	7940	2742	2803
449	<800	3734	3895	5376	5324
2602	2763	1334	1301	1493	1478
2590	2617	2287	2309	823	835
8313	8291	1814	1918	1962	2002
1711	1683	691	<800	2900	2954
9821	9472	5477	5348	1484	1478
516	<800	305	<800	1005	995
8587	8291	25245	25541	1181	1171
7446	7581	3988	4121	18560	19002
2088	2075	1633	1598	3603	3579
2509	2617	631	<800	4943	5076
3519	3501	90	<800	3239	3241
26	<800	402	<800	953	995
925	933	3350	3490	1621	1615
98	<800	4577	4515	8827	6900
6409	6410	1229	1183	79	<800
1383	1376	4221	4331	16418	16263
16048	15829	3980	4121	3361	3241
1876	1877	2548	2497	872	881
4169	4068	674	<800	4220	4193
1674	1683	2332	2309	2848	2803
52	<800	11445	11045	2283	2318
5227	5082	5692	5632	3211	3241
15464	15829	1385	1414	4615	4632
14333	14296	608	<800	6823	6900
1026	1026	3930	3895	886	881
645	<800	1465	1414	2032	2002
6621	6410	2382	2497	8680	8661
8413	8291	747	<800	5163	5076
1401	1376	1951	2012	1547	1478
1512	1471	642	<800	7058	6900
5801	5644	21060	21003	1647	1615
3256	3278	3700	3666	2753	2803
3448	3501	1321	1301	162	<800
2738	2763	10705	10502	1078	1093


```

LOCUS       MMEWRK2               5482 bp    mRNA    linear    ROD 07-JUL-1994
DEFINITION  M. musculus EmrK2 cDNA.
ACCESSION   X78568
VERSION     X78568.1  GI:510664
KEYWORDS    EmrK2 gene; receptor kinase.
SOURCE      Mus musculus.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 5482)
AUTHORS    Choi, K., Wall, C., Hanratty, R. and Keller, G.
TITLE      Isolation of a gene encoding a novel receptor tyrosine kinase from
            differentiated embryonic stem cells
JOURNAL     Oncogene 9 (4), 1261-1266 (1994)
MEDLINE     94181281
PUBMED     8134130
REFERENCE   2 (bases 1 to 5482)
AUTHORS    Choi, K.
TITLE      Direct Submission
JOURNAL     Submitted (05-JUL-1994) Choi K., 1400 Jackson St. #501, Denver Co
            80206, USA
FEATURES    Location/Qualifiers
             1..5482
             /organism="Mus musculus"
             /db_xref="taxon:10090"
             /cell_type="embryonic stem cells"
             249..4241
             /gene="EmrK2"
             249..4241
             /gene="EmrK2"
             /codon_start=1
             /product="embryonic receptor kinase"
             /protein_id="CAA5311.1"
             /db_xref="GI:510665"
             /db_xref="MGD:MG1:95558"
             /db_xref="SWISS-PROT:P35569"
             /translation="MVSCHWOTAVLPVALLGCLLIGVSGSKLVKVPESLKGIVHWQ
             AQGLFLKCRGEAAHNSLPTIVSQEDKSLTIPESCGGRDNQFQSTLILTAQNH
             TGLYCRPLPTSKKKAESSYIFVSDAGSPFHEWTDIPKLVMTIGRLIICRV
             TPNVTITLKEPEDTLPDQRIWDSRGFIANATYKEIGLNCVAGNRLVQIN
             YLTHQTITLDVGRPPSRVLLGQTLVNLCTATLNTVQMSNMVNPVATKRAAS
             IKQIDRSHNNVNFVSVLKINNVESRDGLYICRVKSGSFSQFNTSHVYKAFIS
             VHKRQPOETTAGRRSYRLSKMKVAFSPETVVKQGSATKASRYLHVYSLIKQ
             DVTEADGDTILLGKQSLFNALTLVNVKQIYKASVSLSPPLYPGSKQV
             LUCYVGPRTITLWHPCHHNSKERYDFCTNEESFIDPSSMLGRISISQRM
             VIEGNTLSTLVADSQPGIYSCRAFNKIGIVERNIKFYVDVPGFHWLSLKM
             AEGDKLSCVWNLFYRDTITLRTVNNRTHNSISQNGHSSYSITLNLVKNV
             SLEDSQACARMIYTGEDILKTEVLVROSEAPHLIQNSDYEVSISGTTILCOA
             RGVAPQITLKNKHKIQBPGLIIGPNSLTFIERTVEDEGVRCRSTNKGAVES
             AAYLVQGTSDSKNLELITLTCTCAATLFWLLTFLTRKLSSEVVDLSIIM
             PDEVLDEQERLYDASNEFAERKLKSLGRGAFGVQASAFGKKSPCTRV
             AVKALKESATSEKALMTLEKILTHIGHLNVNLLGACTKGGFLWIVYCYKGN
             LSNYLKSRDLFKDAALHMLKESLEPGLEGQKPLDSVSSSVTSSTFPEDR
             SVSDVEGEDYSEISKQPLTMEDLISYFQVARMFELSKKCIHRDLAARNILSEN
             NVVKICDGLARDIYKNPVYRGGDTIRLPKMAPESTFOKYISKSDVMSYGVILWE
             IFSLGSPYGVQMDQFCSLKEGMRPEYATPEIYQIMLDMWKKPKRERPEAE
             LVEKLGDLIQANVQDQKVIPLNAILTENSNTYPTISERGCGDGFADPHFSGS
             SDQRYVNAFNFMSLERIKTFEELSPNSTSMEDYGLDITLILGSLFLKRFIWTKE
             SKMLDIRIASPKREKGLSDLRPPSPFCSSCGHIRPVQDDESELKESCCSPFDYN
             SWSVLYSSPPA"
BASE COUNT  1464 a 1446 c 1369 g 1203 t

Query Match
Best Local Similarity 31.8%; Score 95.8; DB 10; Length 5482;
Matches 103; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 66 AGACATCGACTTCCTCTGAAATGATGGCTCTGAAATCTATCTTGACAAATCTACAGCA 125
DB 3412 ATATCTGACTTCCTCTGAAATGATGGCTCTGAAATCTATCTTGACAGGCTACAGCA 3471
QY 126 CCAAGACGCGGTGTGCTTACGAGGATATGCTTGCGGAAATCTCTCTTAGG 180
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 3472 CCAATAGCATGIGGICGICATGGCGTCTCTCTGCGGAGATCTCTCTTAGG 3526

RESULT 11
LOCUS     MUSRIK9               6055 bp    mRNA    linear    ROD 09-AUG-1993
DEFINITION Mus musculus receptor tyrosine kinase (FLI) mRNA, complete cds.
ACCESSION U07297
VERSION   U07297.1  GI:293782
KEYWORDS  receptor protein tyrosine kinase.
SOURCE    Mus musculus (strain C57BL/6J, sub-species domesticus); cDNA to
            mRNA.
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 6055)
AUTHORS    Fitzpatrick, H., Kelleher, K., Morris, G.E., Bean, K., Marberg, D.M.,
            Kritz, R., Morris, J.C., Sookdeo, H., Turner, K.J. and Wood, C.R.
            Molecular cloning of murine FLT and FLT4
JOURNAL     Oncogene 8 (8), 2293-2298 (1993)
MEDLINE     9330572
PUBMED     8393164
FEATURES    Location/Qualifiers
             1..6055
             /organism="Mus musculus"
             /strain="C57BL/6J"
             /sub_species="domesticus"
             /db_xref="taxon:10090"
             1..6055
             /gene="FLT"
             27..4028
             /gene="FLT"
             /note="putative"
             /codon_start=1
             /product="receptor tyrosine kinase"
             /protein_id="AAA40078.1"
             /db_xref="GI:293783"
             /translation="MVSCHWOTAVLPVALLGCLLIGVSGSKLVKVPESLKGIVHWQ
             AQGLFLKCRGEAAHNSLPTIVSQEDKSLTIPESCGGRDNQFQSTLILTAQNH
             TGLYCRPLPTSKKKAESSYIFVSDAGSPFHEWTDIPKLVMTIGRLIICRV
             TPNVTITLKEPEDTLPDQRIWDSRGFIANATYKEIGLNCVAGNRLVQIN
             YLTHQTITLDVGRPPSRVLLGQTLVNLCTATLNTVQMSNMVNPVATKRAAS
             IKQIDRSHNNVNFVSVLKINNVESRDGLYICRVKSGSFSQFNTSHVYKAFIS
             VHKRQPOETTAGRRSYRLSKMKVAFSPETVVKQGSATKASRYLHVYSLIKQ
             DVTEADGDTILLGKQSLFNALTLVNVKQIYKASVSLSPPLYPGSKQV
             LUCYVGPRTITLWHPCHHNSKERYDFCTNEESFIDPSSMLGRISISQRM
             VIEGNTLSTLVADSQPGIYSCRAFNKIGIVERNIKFYVDVPGFHWLSLKM
             AEGDKLSCVWNLFYRDTITLRTVNNRTHNSISQNGHSSYSITLNLVKNV
             SLEDSQACARMIYTGEDILKTEVLVROSEAPHLIQNSDYEVSISGTTILCOA
             RGVAPQITLKNKHKIQBPGLIIGPNSLTFIERTVEDEGVRCRSTNKGAVES
             AAYLVQGTSDSKNLELITLTCTCAATLFWLLTFLTRKLSSEVVDLSIIM
             PDEVLDEQERLYDASNEFAERKLKSLGRGAFGVQASAFGKKSPCTRV
             AVKALKESATSEKALMTLEKILTHIGHLNVNLLGACTKGGFLWIVYCYKGN
             LSNYLKSRDLFKDAALHMLKESLEPGLEGQKPLDSVSSSVTSSTFPEDR
             SVSDVEGEDYSEISKQPLTMEDLISYFQVARMFELSKKCIHRDLAARNILSEN
             NVVKICDGLARDIYKNPVYRGGDTIRLPKMAPESTFOKYISKSDVMSYGVILWE
             IFSLGSPYGVQMDQFCSLKEGMRPEYATPEIYQIMLDMWKKPKRERPEAE
             LVEKLGDLIQANVQDQKVIPLNAILTENSNTYPTISERGCGDGFADPHFSGS
             SDQRYVNAFNFMSLERIKTFEELSPNSTSMEDYGLDITLILGSLFLKRFIWTKE
             SKMLDIRIASPKREKGLSDLRPPSPFCSSCGHIRPVQDDESELKESCCSPFDYN
             SWSVLYSSPPA"
BASE COUNT  1633 a 1549 c 1492 g 1381 t

Query Match
Best Local Similarity 31.8%; Score 95.8; DB 10; Length 6055;
Matches 103; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 66 AGACATCGACTTCCTCTGAAATGATGGCTCTGAAATCTATCTTGACAAATCTACAGCA 125
DB 3199 ATATCTGACTTCCTCTGAAATGATGGCTCTGAAATCTATCTTGACAAATCTACAGCA 3258
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

; LENGTH: 5470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 286..4386
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-766-678-1

Query Match 33.8%; Score 710; DB 10; Length 5470;
Best Local Similarity 65.0%; Pred. No. 5.9e-200;
Matches 1081; Conservative 0; Mismatches 575; Indels 2;

QY 201 CACTCTATCTTACCATCATGATGTTTCCCTGCAAGATTCAGGACCATATGCTTGGCAG 260
DB 2148 CATCTTGATGTTGGCATTTCAGATTCCTCTCTGAGGACCAAGGCGACTATTTTGTCTC 2207
QY 261 AGCCAGGAATGTATACACAGGGGAAGAAATCCCTCCAGAGAGAAAGAAATTAACAATCAGAGA 320
DB 2208 TGCTCAAGTANGAGACCCAGAAAGACATTCCTGTCTCAACAGCTCATATCTTAA 2267
QY 321 TCAGGAGACCATACCTCTCCGCAACCTCTAGTATCATCACATGTCGCCATCAGCAGTTC 380
DB 2268 GCGCATGGCACCATGATCACCGGAATCTTGGAGAAATCATCAGACAAACCACTTGGCCAGAC 2327
QY 381 CACCACCTTTAGACTGTATGCTATGATGTCCTCCGAGCTCAGATCACTTGGTTTAA 440
DB 2328 CATTCAGTACTGCTCCCGAGCATCTGGAATCTTACCCACACATATCATGGTTTCAAGA 2387
QY 441 CAACCAAAAATACAAAGAACCTCGGAATATTTTAGGACCAAGAGAGAGCAGCGTGT 500
DB 2388 CAACGAGACCTGGTAGAAGATTCAGGCATTTGACTGAGAGATGGGAACCGGAACCTGAC 2447
QY 501 TATTGAAAGAGTACAGAGAGAGATGAGGTGTCTATCACTTGCACCAAGCCACCAACAGAGA 560
DB 2448 TATCCGACGGGTGAGAGAGAGATGGAGGCGCTTACACCTTGGCAGGCTTGCATGTC 2507
QY 561 GGGCTGTGTGAAAGTTCAGCATACCTCTACTGTTCAGGAACCTCGCACAAAGTCTAATCT 620
DB 2508 TGCTGTGCAAGCGGAGAGCGCTTCTCATATAGAGGTGCCAGGAAAGACCACTT 2567
QY 621 GGAGCTGATCATCTTAACATCAGCTGTGTGGCTCGGACTCTCTCTGGCTTCTATTAAC 680
DB 2568 GGAAGTCATTATCTCTGCTGGCACTGAGTGTGCTATGCTTCTTCTGCTCTCTCT 2627
QY 681 CTCTCTTATCCGAAATGAAAGTCTCTTCTCTGAAATAGAGACTGACTTACCTATC 737
DB 2628 CATGTCTACGGACCGTTAGCGGGGCGCAATGAAGGGGAGCTGAAGACGGCTACTTGTG 2687
QY 738 AATTATTAATGGACCATGATGATGTTCTTTGATGAGCAGTGTGAGGCGGCTCCCTTATGA 797
DB 2688 TATTGTTCATGGATCCAGATGAATTTGCCCTTGGATGAGCGCTGTGAACGCTTGCCTTATGA 2747
QY 798 TGGCAGCAAGTGGAGTTTGGCCCGGAGAGACATTAACCTGGGCAATCACTTGGAGAGAG 857
DB 2748 TGGCAGCAAGTGGAGTTTGGCCCGGAGAGACATTAACCTGGGCAATCACTTGGGCGGCG 2807
QY 858 GGGTTTGGAAAGTGGTTTCAAGCATCAGCATTTGGCATTGAAGAAATCACTTACGTGGCG 917
DB 2808 TGCTTCGGCCAGTGTATGAGGAGAGACGCTTGTGAATTCAGAGACACCGACTTGGCAA 2867
QY 918 GACTTGGGCTGTGAAGTCTGTAAGAGAGGGGCCACGGCCAGCGAGTACAAAGCTGTGAT 977
DB 2868 AACAGTAGCGGTCAAGATGTTCAAGAGAGAGGCAACACACAGCGAGCATCGAGCGCTGAT 2927
QY 978 GACTTGACCTTAAATCTTACCCACATTCGCCACCATCTGAACCTGTGTAACCTGCTGGG 1037
DB 2928 GTCTGAACCTCAAGATCTCTCATCCATTTGTCACCACTTGTCACTTGTGTAACCTGCTGG 2987
QY 1038 AGCCTGCACCAAGCAGGAGGCGCTCTGATGGTGTGATGTTGAATCTGCAAAATATGAAA 1097

```

```

DB 2988 CGCTCGACCAACCGCGGAGGCGCTCTCAIGTGATTTGTGGAATCTGCAAGTTTGGAAA 3047
QY 1098 TCTCTCAACTACTCTCAAGAGCAAAAGTGTACTTATTTTCTCTCAAGATOCACACT 1157
DB 3048 CCTATCACTTACTTACGGGCAAGAGAAATGAATTTGTTCCTTATAGAGCAAAAGGGC 3107
QY 1158 ACACATGGAGCCTAAGABAGAAATATGAGCCAGGCCCTGCGCAACAAAGCAAGCAAG 1217
DB 3108 AGCTTCCCGCAGGCAAGGACTACGTGGG---GAGCTCTCCGAGATCTGAAAAGAGG 3164
QY 1218 ACTAGATAGCTCACCAGCAGCAAGAGCTTTGCGAGCTCCGCGCTTTCAGGAAGATAAAG 1277
DB 3165 CTGGACAGCATCACACAGCAGCAGGCTTCGCCAGCTCAGGCTTCTTGGAGGAGAAATC 3224
QY 1278 TCTAGTGTGTTGAGGAGAGGAGGATCTTACGCTTCTTACAGSAGGCCCATCATCTAT 1337
DB 3225 GCTCAGTGTGTAGAGGAGAGAGAGCTTCTGAAGAAGCTGTACAGGACTTCTGACCTT 3284
QY 1338 GGAAGATCTGATTTCTTACAGTTTCAAGTGGCCAGAGSCATGGAGTTCCTGCTTCCAG 1397
DB 3285 GGAGCATCTCATCTGTTACAGCTTCCAGTGGCTTAGGCAAGGAGTCTTGCAATCAG 3344
QY 1398 AAGTGCATTCATCGGACCTTGGCAGCGAGAGAAATCTTCTTATCTGAGAAACACCTGT 1457
DB 3345 GAATGTATCCACAGGACCTTGGCAGCAGCAACAACTTCTCTATCGGAGAGAAATGTGT 3404
QY 1458 GAAGATTTGTGATTTTGGCTTGGCGGGATTTATTAAGAACCCCATATTTGTGAGAAA 1517
DB 3405 TAGATCTGTGACTTGGCTTGGCGGGGACATTTATTAAGACCCGGAATTTGTGAGAAA 3464
QY 1518 AGGAGATCTCGACTTCTCTGAAATGGATGGCTCCCGAAATCTATCTTTCACAAAATCTA 1577
DB 3465 AGGAGATGCGCGACTCCCTTTGAGTGGTGGTGGCGGCGGAAACCACTTTTCAGAGATA 3524
QY 1578 CAGCAGCAGAGCGAGTGTGGCTTACGGAGTATGCTGTGGGAAATCTTCTCTTAGG 1637
DB 3525 CACAATTCAGACGATGTGGCTTTCGCTTTCGCTGTGGGAAATTTTCTCTTAGG 3584
QY 1638 TGGGCTTCATACCCAGGAGTACAAATGGATGGAGCTTTTGAATCGCTGAGGGAAGG 1697
DB 3585 TGGCTCCCATACCTTGGGTCAGATTAAGAAATTTTGTAGAGATTTGAAGAGG 3644
QY 1598 CATGAGATGAGAGCTCTGTAGTACTTACTCTCTGAAATCTATCAGATCATGTGACTGT 1757
DB 3645 AACTAGAAATGGGCTCTCTGACTACTACTACCCAGAAATGTTACAGACCACTGCTGACTG 3704
QY 1758 CTGCGACAGAGCCCAAGAGAGGCGCAGATTTGCAAGACTTGTGGAAAACACTAGTGA 1817
DB 3705 CTGGCATGAGGACCCCAACAGAGACCCGCTTTTCAGAGTGTGGAGCAATTTGGAAA 3764
QY 1818 TTTGCTTCAAGCAATGTACAAAGAGTGTAAAGACTACAT 1859
DB 3765 CTTCTTSCAAGCAATGCGCAGCAGAGTGGCAAGACTATAT 3806

```

```

RESULT 2
US-09-419-408-5
; Sequence 5, Application US/05919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Iemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Inclone Systems Incorporated
; STREET: 180 Varlick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/919,408
 FILING DATE: 31-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/977,451
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 07/906,397
 FILING DATE: 26-JUN-1992
 APPLICATION NUMBER: US PCT/US92/05401
 FILING DATE: 26-JUN-1992
 APPLICATION NUMBER: TW 81102961
 FILING DATE: 15-APR-1992
 APPLICATION NUMBER: US PCT/US92/02750
 FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593
 FILING DATE: 24-DEC-1991
 APPLICATION NUMBER: US 07/793,065
 FILING DATE: 15-NOV-1991
 APPLICATION NUMBER: US 07/728,913
 FILING DATE: 28-JUN-1991
 APPLICATION NUMBER: US 07/674,666
 FILING DATE: 02-APR-1991

AUTHOR/AGENT INFORMATION:

NAME: Felt, Irving N.
 REGISTRATION NUMBER: 28,601
 REFERENCE/DOCKET NUMBER: LEM 3-79
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-645-1405
 TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 5406 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 FEATURE:

NAME/KEY: CDS

LOCATION: 208..4311

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 265..4308

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 208..264

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-919-408-5

Query Match 33.78; Score 708.4; DB 10; Length 5406;
 Best Local Similarity 55.04; Pred. No. 1.7e-199;
 Matches 1080; Conservative 0; Mismatches 576; Indels 6; Gaps 2;

QY 201 CACTCTTAATCTTACCAATCAATGTTTCCCTGCGAAGATTCAGGCACCTATGCTGCGAG 260

DB 2070 CATCTTGATGGGCAATTCAGATGCTCTCTGCGAGCAAGCGGACATGTTTGTCTC 2129

QY 261 AGCCAGCAATGTATACAGAGGGGAGAAATCTCCAGAGAGAGAAATTAACAATCAGAGA 320

DB 2130 TGCTCAGATAGAGACCAAGAGAGACATTCGCTGGTCAACAGCTCATCTCTAGA 2189

QY 321 TCAGAGAGACCAATACCTCTCGGAAACCTCNGTATACACAGATGGCCATCAGCATTC 380

DB 2190 CGCATGGCCACCAATGATCAGCGGAATCTCGAGATACAGACAAACCATTCGCGAGC 2249

QY 381 CACCACCTTTAGCTGTATGCTATGCTGTCGCCAGGCTCAGATCATCTTGTTTAAAAA 440

DB 3327 TAAGATCTGTGACTTCGGCTTGGCCCGGACATTTATTAAGACCGGATTAATGTGAGAAA 3386

DB 2250 CATTAAGTGAATGCTTGGCCAGCATCTGGAAATCTACCCACACATTAATGTTTCAAGA 2309

QY 441 CAACCCACAAATACAAACAGAGCCTTGAATATTTTATTAGGACAGGAGCAACAGCGTGT 500

DB 2310 CAACGAGACCCCTGGTAGAGATTCAGGCATTTACTGAGAGATGGAGACCGAACCTGAC 2369

QY 501 TATTGAAGAGTGCACAGAGAGAGATGAAGTGTCTATCTACTGCAAGGCCACCAACAGAA 560

DB 2370 TATCCGCGAGGTGAGGAGGAGGATGGAGGCTCTACACCTGCCAGGCTTCAATGTCT 2429

QY 561 GGGCTCTGTGGAAGTTCAGCATACCTCACTGTTCAGGAACCTCGGACAACTCAATCT 620

DB 2430 TGGCTGTGAGAGAGCGGAGACGCTCTCATANTAGAGGTGCGGAGGAGAAAGACCACT 2489

QY 621 GAGCTGATCACTCTAAACATCACCCTGTGTGGTGGCAGCTCTCTTCTGGCTCTTATTAAC 680

DB 2490 GGAAGTCATTAATCTCTGTCGACCTGAGTGGAGTGTTCCTTCTGGCTCTTCTGT 2549

QY 691 CCTCTTAATCCGAAATGAAAGGTC---TCTTCTGAATAAAGACCTGACTACCTATC 737

DB 2550 CATCTGTGACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2609

QY 738 AATTAATGAGCCAGATGAAGTTCCTTTGGATGAGCAGTGTGAGCGGCTCCCTTAIGA 797

DB 2610 TATGTCAIGSATCCAGATGATGCTTGGATGAGCGCTGTGAGCGCTTGGCTTAIGA 2669

QY 798 TCCAGCAAGTGGGAGTTTGGCCGGGAGACACTTAAACTGGGCAATCCTTTGGAAGAG 857

DB 2670 TCCAGCAAGTGGGAGTTTCCCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2729

QY 858 GGCCTTTGGAAGTGGTTCAGCATCAGCAATGCGATTTGCAATTAAGCAATCACCCTAC 917

DB 2730 TGCCTTCGGCCAAAGTATTCAGGCGAGCGCTTTTGAATTTGACAAGACGCGACT 2789

QY 918 GACTGTGCTGTGAAAATGCTGAAAGAGGGGGCGGCGGCGGAGTACAAAGCTGTAT 977

DB 2750 AACAGTAGCGTCAAGATTTGAAGAGAGGAGCAACACACAGCAICGAGCGCTCAT 2849

QY 978 GACTGAGCTAATAATCTTGACCCACATTTGCCACCATCTGAACGTGTTTACCTGTGGG 1037

DB 2850 GTCGAACTCAAGATCTCTCCATTTGTCACCATCTCAATGTTGTTGAGCTCTCTAG 2909

QY 1038 AGCTGTGACCAAGCAAGAGGAGGCGCTCTGATGTTGTTGTAATCTGCAATATGNA 1097

DB 2910 CGCTGTGACCAAGCGGAGGCGCTCTGATGTTGTTGTAATCTGCAATATGNA 2969

QY 1098 TCTCTCACTACTCTCAAGAGCAACGCTGACTTATTTTCTCAACAGAGATCGACACT 1157

DB 2970 CCTATCAACTTACTTACGGGCGAAGAGAAATGAATTTGTTCCCTATAGAGCAAGGG 3029

QY 1158 ACATGAGGAGCTTAAGAGAGAAATGGAGCGGCGCTGGAGAGGAGGAGGAGGAGGAG 1217

DB 3030 AGCTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3086

QY 1218 ACTAGATAGGCTCACCAGCAGCAAGAAAGCTTTGGAGCTTCGGCTTCAGGAGAAATAAG 1277

DB 3087 CTTCGAGCAGTATCACCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3145

QY 1278 TCTGAGTGTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1337

DB 3147 GCTCAGTGTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3206

QY 1338 GGAAGATCTGATTTCTACAGTTTTCAGTGGCCAGGAGGAGGAGGAGGAGGAGGAGGAG 1397

DB 3207 GAGCATCTCATCTGTACAGCTTCCAGTGGCTTAAGGCAATGGAGTTTCTTGGCATCAG 3266

QY 1398 AAGTGTCAATTCAGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1457

DB 3267 GAAGTGATCCACAGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3326

QY 1458 GAAGATTTGATTTTGGCTTGGCCGGGATATTTATAAGAACCCCGGATTAATGTGAGAA 1517

DB 3327 TAAGATCTGTGACTTCGGCTTGGCCCGGACATTTATTAAGACCGGATTAATGTGAGAA 3386

SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-982-610-45

Query Match

Best Local Similarity 25.84; Score 541.6; DB 10; Length 3108;
Matches 981; Conservative 0; Mismatches 679; Indels 12; Gaps 2;

QY 185 CTAGGAGGACCTCCATCAGCTTAACTGATACCAATGATGATGTTTCCCTGCNAGATCAG 244
DB 2848 CTGGGCGGCGGACCGGCTGAGCTGATATCCCGGCGTCCGCGCGGACGAGG 2907
QY 245 GCACCTAGCTTCAGAGCCAGGAAATGATACAGAGGAGGAAATCTCCAGCAAGAG 304
DB 2908 CCACTAATGCTGGAAGTGAAGACCGCGAGCCATGACAGACACTGCGCAAGAGT 2967
QY 305 AATTAATACAGATCAGAGACACCTACTCTCTGGAACTCTGATGATCAGACAG 364
DB 2968 ACCTGTGGGTGAGGCGCTGAGACCCCTGGCTCAGCAGAACTTGACGAGCTCTGG 3027
QY 365 TGGCATCAGAGTTCACACCACTTACACTGTCATGCTAATGGTGTCCGAGGCTCAGA 424
DB 3028 TGAAGTGAAGGACTCGCTGGAGATGAGTCTGTGGCGGAGCGGACCGCCAGCA 3087
QY 425 TCAGTTGGTTTAAACACACCAAAATACACAGAGCTTGGAAATTAATTAAGACAG 484
DB 3088 TCGTGTGTAAGAGCAGAGGCTGCTGGAGAAAGTCTGGAGTGCAGCTTGCGACT 3147
QY 485 GAACGACGCTGTTTATGAAGAGTACACAGAGAGATGAAGGTGCTATCAGTGA 544
DB 3148 CCAACAGAGCTGAGATCAGGCGCTGGCGAGGAGTGGGAGCGGCTATCTGTCGA 3207
QY 545 AAGCCACCAACGAGGCTCTGTGAAAGTTACAGTACCTCACTGTTCAAGAGACT 604
DB 3208 GCGTGTCAACGCCAGGCTCGCTCACTCTCCGCCAGCGTGGCGGTGAAGGCTCGG 3267
QY 605 CGNCAAGTCTAATCTGAGCTGATCACTTACATGCACTGTGTGCTGCGACTCTCT 664
DB 3268 AGATTAAGGCGAGCTGAGATGAGATGCTGATCTTGTGCTACCGGCTCATGCTCTCT 3327
QY 665 TCTGGTCTCTTAACTCCAGCTATCGGAAATGAAGAGTCTTC---TCTCAATAA 721
DB 3328 TCTGGTCT 3387
QY 722 AGACTGACTACTCAATTAATGAGCCAGGATGAGTCTCTCTCTCTCTCTCTCTCTCTCT 781
DB 3388 AGAGGCTCTACCTGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3447
QY 782 ACCGCTCTCTTATCATCCAGCAGTGGGATTTGCCGCGGAGAGACTTAACTGGCA 841
DB 3448 AATACCTGTCTTACATGCCAGCAGTGGGAAATCCCGGAGCGGCTGACCTGGGA 3507
QY 842 AATCACTTGAAGAGGCTTTTGGAAAGTGGTTCAAGCATCAGCATTTGGCAITAGA 901
DB 3508 GAGTGTCTGGCTACGGCGCTTGGGAAGTGGTGGAGCTCTGCTTTCGCAITACA 3567
QY 902 AATCACTACGTGCGGACTGTGGCTGTGAAATGCTGAAGAGGGGCGGCGGCGGAGG 961
DB 3568 AGGGAGCAGCTGTGACACCGCTGGCCCTGAAATGCTGAAAGAGGGGCGGCGGCGGAGG 3627
QY 962 AGTACAAAGCTGTGAGTACTGAGTAAATCTTGACCAATTTGGCCACCATCTGAGCG 1021
DB 3628 AGACCGCGGCTGAGTCTGGAGCTCAAGATCTTATCATGCTGCAACCATCTCAAG 3687
QY 1022 TGTTAACTGTGGAGCTCTGACCAAGCAGAGGCGCTCTGATGGTGAITGTTGAAT 1081
DB 3688 TGGTCAACCTCTCTGGGCGGTGACCAAGCGGCGGCGGCTCTGATGGTGAITGTTGAAT 3747
QY 1082 ACTGCAATATGAATCTCTCACTCTCACTCTCACTCTCACTCTCACTCTCACTCTCACT 1141
DB 3748 TCTGCAAGTACGCAACCTCTCTCACTCTCTCACTCTCTCACTCTCTCACTCTCTCACT 3807
QY 1142 AAGAGTGCAGCTACATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1201

DB 3808 GCGCGAGAGATCTCCGA-----GCAGCGGAGAGCTTCCGCGCCATGCTGGAGC 3858
QY 1202 AAGCAAGAAACCAAGACTAGATAGCTGACAGAGCGGAGAGCTTTCGAGCTCCGGCT 1261
DB 3859 TCGCAGGCTGATCGGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3918
QY 1262 TCGAG 1321
DB 3919 CGAAGAGCGGAG 3978
QY 1322 AGAGAGCCATACATGGAAGATCTGATTTTACAGTTTTCAGTGGCCAGAGGATGG 1381
DB 3979 TGAGCGCTGAGCAAG 4038
QY 1382 AGTTCCTGCTTCCAGAAAGTGCATTCAGGAGCTGCGAGAGAGAGAGAGAGAGAGAG 1441
DB 4039 AGTTCCTGCTTCCAGAAAGTGCATTCAGGAGCTGCGAGAGAGAGAGAGAGAGAGAG 4098
QY 1442 CTGAG 1501
DB 4099 CGAAGAGCGAGCTGAGAGATCTGATTTTGGCTTGGCCGAGAGAGAGAGAGAGAGAG 4158
QY 1502 CCGATATGTGAGAAAGAGAGATCTGAGTCTCTGAAATGAGTGGCTCCGAGATCTA 1561
DB 4159 CTGAGTGTGCGAG 4218
QY 1562 TCTTGACAAATCTACAGCAG 1621
DB 4219 TCTTGACAAAGTGTACAGCAG 4278
QY 1522 AATCTCTCTTCTGAGTGGTCTCCATACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
DB 4279 AGATCTCTCTCTGAGGAGCTTCCCGGTACCTGGGTGAGATCAATGAGAGAGAGAGAGAG 4338
QY 1522 GTGCGCTGAGGAG 1701
DB 4339 AGCGGTGAG 4398
QY 1742 AGATCTCTGAGTCTGCGCAG 1801
DB 4399 GATCTGCTGAGTCTGCGCAG 4458
QY 1802 TGGAAAG 1853
DB 4459 TGGAGATCTGAGGAG 4510

RESULT 6

US-09-982-610-17/c
: Sequence 17, Application US/09982610
: Patent No. US2002014620A1

GENERAL INFORMATION:

: APPLICANT: Genentech, Inc.
: Bennett, Brian D.
: Goeddel, David
: Lee, James M.
: Matthews, William
: Tsai, Siao Ping
: Wood, William J.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080

COMPUTER READABLE FORM:

: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WinPatIn (Genentech)

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/982,610
: FILING DATE: 17-Oct-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/446,648
: FILING DATE: 1996-MAY-23
: APPLICATION NUMBER: 08/222616
: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M.
: REGISTRATION NUMBER: 40,378
: REFERENCE/DOCKET NUMBER: P0821P3PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6827 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-982-610-17

Query Match 13.5%; Score 282.8; DB 10; Length 6827;
Best Local Similarity 59.2%; Pred. No. 3.5e-73;
Matches 386; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 1296 AGAGAGATCTCGCGGTTCTACAGAGCCGACATGATGGAATCTGATTTTGA 1355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2783 AGACCAAGAAGCTGAGGACCTGTGCTGAGCCGCTGACCATGGAAGATCTGTCGTA 2724

QY 1356 CAGTTTCAAGTGGCAGAGGATGGATCTCTGCTTCACAGAAATGCAATCATCGGGA 1415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2723 CAGTTTCAAGTGGCAGAGGATGGATCTCTGCTTCACAGAAATGCAATCATCGGGA 2564

QY 1416 CCTGCCAGGAGAAATCTTTTATCTGAGAACCAAGCTGGTGAAGATTTGATTTTGG 1475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2663 CCTGGCTGCTCGGAATCTCTGCTGCGAAGGAGGCTGGTGAAGATCTGATTTTGG 2604

QY 1476 CTTTCCCGGGATTTATAGAACCCCGATTATGTGAGAAAGAGATATCGACTTCC 1535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2603 CTTTCCCGGGATTTATAGAACCCCGATTATGTGAGAAAGAGATATCGACTTCC 2544

QY 1536 TGTAAATGATGGCTCCCGAATCTATCTTTGACAAAATCTACAGCACCAAGAGCGAGT 1595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2543 CTGAGATGATGGCCCTGAGAGCATCTTCGACAGGCTGACACAGGAGAGTGAAGT 2484

QY 1596 GTGGTCTTACGAGATCTCTGTGGAAATCTTCTCTTAGTGGTCTCCATACCCAGG 1655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2483 GTGGTCTTACGAGATCTCTGTGGAAATCTTCTCTTAGTGGTCTCCATACCCAGG 2424

QY 1656 AGTACAAATGGATGAGGACTTTTGAGTTCGCTGAGGAGGCAATGAGATGAGAGTCC 1715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2423 GGTGCAGATCAATGAGGAGTCTGCCAGCGCTGAGAGAGGCAAGATGAGGCGCCC 2364

QY 1716 TGAGTACTTACTCTGAAATCTATCAGATCATCTGCTGGACTCTGCGACAGACCCAAA 1775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2363 GGAAGTGGCCACTCCCGGCAATAGCGGCAATCATGCTGAATCTGCTCGGAGACCCAA 2304

QY 1776 AGAAGGCCAAGATTTGAGAACATTTGGAAACACTAGGATTTGTTTCAAGCAATGT 1835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2303 GCGAGAGCTTGCATCTCGGAGCTGGTGGAGATCTTGGGGGACCTGCTCCAGGCGAGG 2244

QY 1836 ACAACAGATGTTAAAGA 1853
    ||| ||| ||| |||
DB 2243 CTGCAAGAGAGAGGA 2226
    ||| ||| ||| |||

RESULT 7
US-09-954-531-1383

: Sequence 1383, Application US/09954531
: Patent No. US20020165180A1
: GENERAL INFORMATION:
: APPLICANT: Weaver, Zoe
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
: FILE REFERENCES: 589290-77
: CURRENT APPLICATION NUMBER: US/09/954,531
: PRIOR FILING DATE: 2002-05-02
: PRIOR APPLICATION NUMBER: US/60/233,133
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,009
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,034
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,509
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: US/60/234,567
: NUMBER OF SEQ ID NOS: 1392
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1383
: LENGTH: 5084
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-954-531-1383

Query Match 12.3%; Score 259; DB 9; Length 5084;
Best Local Similarity 55.8%; Pred. No. 3.4e-66;
Matches 595; Conservative 0; Mismatches 440; Indels 32; Gaps 4;

QY 790 CTTATGATGCCAGCAGTGGAGATTTGCCCGGAGAGAGACTTAACTGGGCAATCACTT 849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1747 CTTCTTATGATCACAATTTGGAGATTTCCAGAAACAGGCTGAGTTTGGGAAACCCCTG 1806

QY 850 GGAAGAGGGCTTTTGGAAAAGTGGTTCAAGCATCAGCATTTGGCAATTAAGAAATCACT 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1807 GGTGCTGAGCTTTTGGGAGGTTGTGAGCAGACTGCTTATGGCTTAATTAAGTCAGAT 1866

QY 910 AGTTCGCGGACTGTGGCTGTGAAATGCTGAAGAGGGGGCCACGCCAGCGAGTACAAA 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1867 GCGGCCATGACTGTGCTGTAAAGAAGTCAAGCCGAGTGGCCATTTGACAGAACGGAA 1926

QY 970 GCTCTGATGACTGAGCTAATAATCTTGACCACATTTGGCCACCATCTGACGTGGTTAC 1029
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1927 GCGCTCATCTCTCACTCAAGCTCTGAGTTTACCTGGTAATCATGATATTTGTAAT 1986

QY 1030 CTGCTGGAGCCTTGACCAAGCAAGAGGAGGCTCTGATGGTGAATTTGAAATACGCAAA 1089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1987 CTACTTGGAGCCTTGCACCA- --TTGGAGGCCCCACCTGGTCATTTACAGATATTTGTC 2043

QY 1090 TATGGAATCTCTCCAATCTACCTCAAGAGCAAGAGTACTTATTTTTCACAAG- -- 1146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2044 TATGGTGAATCTTTGAATTTTGGAGAGAAAGAGTGAATTTTATTTTGTTCAGAGCAG 2103

QY 1147 -----GATCGAGCACTACACATGAGCCCTAAGAAAGAAAATGGAGCCAGCG 1194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2104 CAAGATCATGCAGAGCTGCATTTATAAGAATCTTCTCATTTCAAGAGGATCTCTG- 2162

QY 1195 CTGGAACAAGGCAAGAAACCAAGACTAGTAGCGTCACCAAGAGGAGGCTTTGGAGC 1254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2163 CAGCGATAGTACTAATGATGATGACATGAAACCTGGAGTTTCTTAIGTTGTCCCAAC 2222

QY 1255 TCCGGCTTTTCAGGAAGATAAAAGCTGAGTGTGTTTGGAGAGAGGAGGATCTCTGAGGT 1314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2223 CAGGCCGCAAAAAGAGATCTGTGAGNATAGGCTCATACATAGAAAGNATGAGTCTC 2282

QY 1315 TTT-----TACAAGGAGCCCATCACTATGGAAGATCTGATTTTCTTACAGTTT 1361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2283 CCGCATCATGAGGAGTACGAGTTGGCCCTAGACTTAGAAGACTTGTGAGCTTTTCTTA 2342

QY 1362 TCAAGTGGCCAGAGGCAATGGAGTCTCTGCTTCCAGAAAGTCAATTCATGGGACCTGGC 1421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

D	b	2343	C	A	G	G	T	G	G	C	A	A	G	G	C	A	T	G	G	C	T	T	C	T	C	C	C	T	C	C	A	A	A	T	T	G	T	A	T	T	C	A	C	A	G	A	G	A	C	T	T	G	G	C		2402
O	y	1422	A	G	C	A	G	A	A	A	C	A	T	C	T	T	I	A	T	C	T	G	A	A	C	A	G	T	G	T	G	A	A	T	T	G	A	T	T	G	A	T	T	T	G	G	C	T	G	C		1481				
D	b	2403	A	G	C	A	G	A	A	T	A	T	C	C	T	T	I	A	T	C	A	T	C	A	G	T	C	G	A	T	C	A	A	N	A	T	T	G	A	T	T	T	T	T	T	T	T		2452							
O	y	1482	C	G	G	G	A	T	A	T	A	A	G	A	C	C	C	G	A	T	A	T	G	T	G	A	A	A	G	A	A	T	A	T	G	A	A	A	G	A	T	C	C	T	C	T	G	A		1541						
D	b	2463	C	A	G	A	G	A	C	A	C	A	G	A	T	A	T	T	A	T	T	G	T	G	T	T	A	A	G	A	A	C	C	T	G	A	C	T	A	C	T	G	T	G	A		2522									
O	y	1542	A	T	G	T	G	G	T	C	C	G	A	A	T	A	T	C	T	T	G	A	A	A	T	C	A	C	C	C	A	A	T	C	A	C	C	A	G	A	G	C	G	T	G	T	C		1601							
D	b	2523	T	G	S	G	T	G	C	C	A	C	T	T	C	A	T	T	C	A	C	T	T	G	A	C	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		2582								
O	y	1602	T	A	C	G	A	G	A	T	T	C	T	T	G	G	A	A	T	C	T	C	T	T	A	G	T	G	G	H	T	C	C	A	T	A	C	C	C	A	G	A	G	A	T	A		1661								
D	b	2583	C	T	A	T	G	G	A	T	T	C	T	T	T	G	G	A	T	G	T	C	T	T	T	A	G	A	G	C	C	C	T	A	T	C	T	T	T	T	T	T	T	T	T		2642									
O	y	1662	R	A	T	G	A	T	G	A	G	C	A	T	T	T	T	C	A	T	T	T	C	A	T	T	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		1721								
D	b	2643	G	E	T	C	A	T	T	C	A	G	T	T	C	A	A	G	A	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		2702									
O	y	1722	C	T	C	A	T	C	T	T	G	A	A	T	C	T	A	C	A	T	C	T	G	A	T	C	T	G	A	T	C	T	G	A	T	C	T	G	A	T	C	T	G	A	T	C	T		1781							
D	b	2703	C	C	A	C	T	C	T	G	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		2762									
O	y	1782	C	C	M	A	G	A	T	T	G	C	A	A	T	T	C	T	G	A	A	A	A	A	C	T	A	G	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		1842									
D	b	2763	A	C	C	A	A	T	T	C	A	A	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		2903									

RESULT 8

- : US-09-967-768A-277
- : Sequence 277, Application US/09967768A
- : Patent No. US20020150877A1
- : GENERAL INFORMATION:
- : APPLICANT : Augustus, Meena
- : TITLE OF INVENTION : Cancer Gene Determination and Therapeutic Screening
- : TITLE OF INVENTION : Sels
- : FILE REFERENCE : 68929

```

: RESULT 9
: US-09-919-408-3
: Sequence 3, Application US/09919408
: Patent No. US20020072077A1
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: TOTOPIENT HEMATOPOIETIC STEM CELLS
: RECEPTORS AND THEIR LIGANDS
:
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ImClone Systems Incorporated
: STREET: 180 Varick Street
: CITY: New York

```

STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10014
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/919,408
 FILING DATE: 31-Jul-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/977,451
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 07/906,397
 FILING DATE: 26-JUN-1992
 APPLICATION NUMBER: US PCT/US92/05401
 FILING DATE: 26-JUN-1992
 APPLICATION NUMBER: IW 81102961
 FILING DATE: 15-APR-1992
 APPLICATION NUMBER: US PCT/US92/02750
 FILING DATE: 02-APR-1992
 APPLICATION NUMBER: US 07/813,593
 FILING DATE: 24-DEC-1991
 APPLICATION NUMBER: US 07/793,065
 FILING DATE: 15-NOV-1991
 APPLICATION NUMBER: US 07/728,913
 FILING DATE: 28-JUN-1991
 APPLICATION NUMBER: US 07/679,666
 FILING DATE: 02-APR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Feit, Irving N.
 REGISTRATION NUMBER: 28,601
 REFERENCE/DOCKET NUMBER: IEM-3-7P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-645-1405
 TELEFAX: 212-645-2054
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3501 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 58..3039
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 139..3035
 FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: 58..138
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-09-919-408-3

Query Match 11.9%; Score 249.4; DB 10; Length 3501;
 Best Local Similarity 55.9%; Pred. No. 1.9e-63;
 Matches 572; Conservative 0; Mismatches 421; Indels 30; Gaps 4;

QY 792 TATGATGCCAGCAAGTGGAGTTGGCCGGGAGAGNCITAACTGGGCAAACTCTGG 851
 Db 1846 TGAATATGATCTCAAAATGGAGTTTCCAGAGAGAAATTTAGAGTTGGCAAGTACTAGG 1907
 QY 852 AAGAGGGGCTTTTGGAAAGGTTTCAAGCAATCAGCAATTTGGCATTAAGAAATCACCTAC 911
 Db 1908 ATCAGGTGCTTTTGGAAAGGATGAGCGGACGCTTAAGGAAATGCAAGAGGAGT 1967
 QY 912 GTGCCGCTGTGGCTGTGAAAATGCTTGAAGAGGGGGGCGCCAGCGAGTACAAAGC 971

Db 1968 CTCAATCCAGGTTGCCGTCAAAAATGCTTGAAGAAAGAGCAGACAGCTCTGAAAGAGAGGC 2027
 QY 972 TCTGATGACTGAGCTAAATAATCTTGACCCCAATGGCCCAATCTTACAGCTGGTTTACCT 1031
 Db 2028 ACTCATGTGAGACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTTGTAACCT 2087
 QY 1032 CTCTGGAGCCTGCACCAGCAAGAGGGGCTCTGATGGTATTTGATTTGATTAACAATA 1091
 Db 2088 CTCTGGGGGCGACACACTGTGAGG--CCAATTTACTGTTTGGATTTGATTTGCTA 2144
 QY 1092 TGGAAATCTTCAACTACCTCAAGCAAAAGCTG-----ACT 1129
 Db 2145 TGGTATCTTCTCAACTATTAAGAAGTAAAGAGAAATAATTCACAGGACTTGGACAGA 2204
 QY 1130 TATTTTCTCAACAAGGATGACGACACTACACATGAGGCTTAAGAAAGAAAAATGGAGC 1189
 Db 2205 GATTTTCAAGGAACAAATTTTACGTTTACCCCACTTTTCCATCACAATCCAAATTCAG 2264
 QY 1190 CAGGCTTGGAAACCAAGAACCAAGACTAGATAGCGTACACAGAGGGAAGCTT-- 1247
 Db 2265 CATGCTTGGTTCAGAGAAAGTTTCAATACACCGGACTCGGATCAAAATCTCAGGGCTTCA 2324
 QY 1248 TCCGAGCTCCGGCTTTTCAAGAAAGATAAAAGTCTGAGTGATGTTGAGGAAGAGGAGATTC 1307
 Db 2325 TCGGAATTCATTTTCACTCTGAGATGAATTTGAATATGAACCAAAAGAGCTTGAAGA 2384
 QY 1308 TGACGCTTTTCAAGAGGCGCCATCACTATGGAAGATCTGATTTCTTACAGTTTCAAGT 1367
 Db 2385 AGAGGAGGCTTCAATGTGC---TTACATTTGAAGATCTTCTTTCCTTGCATATCAAGT 2441
 QY 1368 GGCAGAGCAATGAGTCTGCTTCCAGAAAGTCAATTCATCGGGACCTGCGAGGAG 1427
 Db 2442 TGCCAAAGGAATGGAATTTCTGGAAATTTAAGTCTGTGTTACAGAGACCTTGCSCCAG 2501
 QY 1428 AAACATTTCTTTATCTGAGAACACAGCTGTGAAGATTTGTGATTTTGGCTTGCCTGGGA 1487
 Db 2502 GAAGCTGCTTTCACCCCGGAAAGTGGTGAAGATATGTGATTTGGATTGGCTCGAGA 2561
 QY 1488 TATTTAAGACCCCGGATTTGTGAGAAAGGAGATCTCGACTTCTTGAATGGAT 1547
 Db 2562 TATCATGATGATTCCAACTATTTGTGAGGGCAATGCCCTCTGCTCTAAATGGAT 2621
 QY 1548 GGCTCCGGAAATCTTTTGAACAAATCTACAGCACCAAGAGCGGCTGTGCTTACGG 1607
 Db 2622 GGCCCGGAAAGCTTTTGAAGGCACTTACACCATTAAGATGATGTTGGTCAATGG 2681
 QY 1608 AGTATCTGTGGAAATCTTCTCTTAGGTGGTCTCCATACCCAGAGTACAAATGGA 1667
 Db 2682 AATATTAATCTGTGGAAATCTTCTCACTTGGTGAATCTTACCTTGGCATTCGGGTGA 2741
 QY 1668 TGAAGGCTTTTGCAGTCCCTGAGGAGGCAATGAGGATGAGGCTCTCTGATCTTAC 1727
 Db 2742 TGTAACTTCTCAAACTGATTCAAATGGATTTAAATGGATTAAGCATTTTATGCTAC 2801
 QY 1728 TCTTGAATATCAATGATGCTTGGACTCTGTCACAGAGACCCCAAGAAAGAGGCAAG 1797
 Db 2802 AGAAGAAATATCAATTAATGCAATCTCTGCGGCTTTTGGACTCAAGGAACGGCCATC 2861
 QY 1788 ATT 1790
 Db 2862 CTT 2864

RESULT 10
 US-09-872-136-3
 : Sequence 3, Application US/09872136
 : Patent No. US2002019545A1
 : GENERAL INFORMATION:
 : APPLICANT: Lemischka, Ibor R.
 : TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
 : RECEPTORS AND THEIR LIGANDS
 : NUMBER OF SEQUENCES: 10

TYPE: DNA
ORGANISM: Homo sapiens
US-09-944-807-9

Query Match 11.1%; Score 233.6; DB 10; Length 3992;
Best Local Similarity: 54.3%; Pred. No. 1e-58;
Matches 539; Conservative 0; Mismatches 444; Indels 9; Gaps 3;

QY 805 AAGTGGGAGTTCCTGGGAGAGACATTAACCTGGGAAATACATTCGAGAGGGGCTTTT 864
|||||
DB 2020 AAGTGGGAGTTCCTGGGAGAGACATTAACCTGGGAAATACATTCGAGAGGGGCTTTT 2079
|||||

QY 865 GGAAAGTGGTTCAGCATCAGCATTTGGCATTAAAGAAATACCTACGTGCGGACTGTG 924
|||||
DB 2080 GGAAGGTGGTTCAGCATCAGCATTTGGCATTAAAGAAATACCTACGTGCGGACTGTG 2139
|||||

QY 925 CTTGTAAATCTGAAAGAGGGGCGCAGCGGAGAGATCAAGAGCTCTGATGACTGAG 984
|||||
DB 2140 CTTGTAAATCTGAAAGAGGGGCGCAGCGGAGAGATCAAGAGCTCTGATGACTGAG 2199
|||||

QY 985 CTAAATCTTGACCTACATTCGACCATCTGACCTGAGCTGTTAACTCTGAGAGCTGC 1044
|||||
DB 2200 CTGAGATCAATGAGCTACCTGCGGAGGAGAGAGATCTGCTGAGAGCTGT 2259
|||||

QY 1045 ACCAAGCAGAGGCGCTCTGATGATGTTTCAATAGTCAAAATATCGAAATCTCTCC 1104
|||||
DB 2260 ACCAAGCAGAGGCGCTCTGATGATGTTTCAATAGTCAAAATATCGAAATCTCTCC 2316
|||||

QY 1105 AACTACTCTCAAGAGCAAGGTGA---CTTATTTTCTCAAGAGATGAGAGCTACAC 1161
|||||
DB 2317 AACTTTCTGGAAGAGAGGCTGAGGCTATGCTGGGACCGAGCTGAGCGCGGCGAGGAC 2375
|||||

QY 1162 ATGAGGCTTAAGAAAGAAATAGAGCGAGGCTTGAACAGAGGCAAGAAACCAAGACTA 1221
|||||
DB 2377 CCGGAGGAGGCTGTGACTATAAGAAATCAACCTCGAGAGAAATATGTCGAGGAGGAC 2436
|||||

QY 1222 CATAGCTCATCAGAGAGGAGAGCTTTGGAGGCTCCGGCT---TCAGGAAGATAAAGT 1278
|||||
DB 2437 AGTGGCTCTCCAGAGGCTGTGAGACACCTATGTGAGATGAGGCTGTCTCCTCTCT 2495
|||||

QY 1279 CTGAGTGTATGAGAGAGAGAGATCTGACGGTTTCTCAAGAGGAGGCTCACTATG 1338
|||||
DB 2497 TCAATGACTCTCTGACAGAGAGCTGAGCAAGGAGATGGAGCGGCTCTGGAGCTC 2556
|||||

QY 1339 GAAGATCTGATTTCTACAGTTTCAAGTGGCCAGAGAGGATGGAGTCTCTGTTCTCAGA 1398
|||||
DB 2557 CGGAGCTCTCTACCTCTCCAGCAAGTAGCCAGGAGATGGCTCTCTGCTCTCTCAG 2616
|||||

QY 1399 AAGTGCATTCAGGAGCTGCGAGGAGAGAAATCTTTTATCTGAGAGAGAGCTGGTG 1458
|||||
DB 2617 ANTGTATCCAGGAGGAGTGGAGGAGCGTAGCGTGTGAGTGTGAGTGTGAGGCT 2676
|||||

QY 1459 AAGATTGTGATTTGGCTTCCCGGGATATTAAGAACCCGATTAATGTGAGAAAT 1518
|||||
DB 2677 AAGATTGGGAGCTCGGCTGCTAGGAGCATCATGATGACTCCAGTCACTATGTCAAG 2736
|||||

QY 1519 GGAGATCTGAGCTCTCTGAAATGAGATGGTCCGAACTATCTTTTGACAAATCTAC 1578
|||||
DB 2737 GGCAATGCCCGGCTCTGCTGAGATGAGTGGCCAGAGAGCATTTGACGTGCTAC 2796
|||||

QY 1579 AGCAACAGAGGAGGAGTGTGCTTACGAGTATTCGTGTGGAAATCTCTCTTAGGT 1638
|||||
DB 2797 ACGGTTCAGAGGAGGCTGTGCTTACGAGTATTCGTGTGGAAATCTCTCTTAGGT 2856
|||||

QY 1639 GGGTCTCATACCGAGGAGTACAAATGATGAGGAGCTTTGAGTGGCTGAGGAGGAGC 1698
|||||
DB 2857 CTGAATCCCTACCTGGGATCTCTGGTGAAGAGAGAGTCTATTAAGTGGTGAAGAGTGA 2916
|||||

QY 1699 ATGAGGATGAGCTCTGAGTCTCTGATCTCTGAAATCTATCAATCTGCTGAGTCTG 1758
|||||
DB 2917 TACCAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2975
|||||

QY 1759 TGGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1790
|||||

Db 2977 TGGGCTTGGAGCCACCCACAGACACCCACCTT 3008
|||||

RESUL 13
US-09-919-408-1
Sequence 1. Application US/09919408
Patent No. US20020072077A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/919,408
FILING DATE: 31-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,451
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 112..3006
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 31..111
FEATURE:
NAME/KEY: CDS

LOCATION: 31..3009
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-919-408-1

Query Match 10.5%; Score 221.6; DB 10; Length 3453;
Best Local Similarity 53.0%; Pred. No. 3.3e-55;
Matches 565; Conservative 0; Mismatches 489; Indels 30; Gaps 4;

QY 785 GGCCTCCCTATATGATCCAGCAAGTGGAGTTTGCCCGGAGAGACITAAATGCGCAAA 844
DB 1817 GGGACTATGAATAICACTTAAGTGGAGTTTCCCGAGAGAACTTAGAGTTTGGGAAG 1876
QY 845 CACITGGAAGAGGCGCTTTTGAAGAAGTGTTCAGCAACAGCAATTTGCAATTAAGAAAT 904
DB 1877 TCTGGGCTGCGCTTTTGGAGGGGTATGAACCCAGCGGCTTAGCATTATGAATAA 1936
QY 905 CACCTACGTGCGGAGCTTGCGTGTGAATAATGCTCAAGAGGGGCGCCAGCGCCAGCAGT 964
DB 1937 CGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTCTAAGAGAAAGCTGACACGCTGGAA 1996
QY 965 ACACAGCTCTGATGACTGACCTAATAATCTTGACCCACATTTGCCACATCTGAACGTGG 1024
DB 1997 AGAGAGCTCTGATGCGGCTCAAAATGATGACCCACCTGGGACACCAATGACAAATCG 2056
QY 1025 TTAACCTCTGGGAGCGCTGCAACCAAGAGAGGCGCTCTGATGTTGTTGATTAACI 1084
DB 2057 TGAATCTGCTGGGCAAGCAC---ACTGTAGGGCCAGTGTACTTGATTTTGAATAT 2113
QY 1085 GCAATATGGAATCTTCCACTACTCTCAAGAGCAAAACGTT----- 1126
DB 2114 GTTGCTATGGTGAATCTTCACTACTAAGAGTAAAGAGAGAAAGTTTACAGGACAT 2173
QY 1127 ---ACTTATTTTCTCAACAGGATGACGACTACAGTACAGGAGCTTAAGAAAGAAAA 1182
DB 2174 GGACAGAGATTTTAAGGAACATAATTTCACTGTTCTTACCTCTTCCAGGACAAATCAA 2233
QY 1183 ATGGAGCCAGCGCTGGAACAGGAGCAAGAAACCAAGACTAGATAGCGTACACAGAGCGAA 1242
DB 2234 ATTCCAGCATGCGTGTGTACAGAGAGTTCAGTTACACCCGCTTGGATCAGCTCAG 2293
QY 1243 AG---CTTTCCAGCTCGCGTTTCAGGAGATAAAGTCTGAGTGTGTTGAGGAAGA-- 1298
DB 2294 GGTCAATGGGAATTAATTAATCTGAGAGATGAGATTAAGTAAGAAACCAAGAGAGGC 2353
QY 1299 -GGAGGATCTGCGGTTTCTACAGAGGAGCCATCACTATGGAAGATCTGATTTCTTACA 1357
DB 2354 TGGCAGAAGAGAGGAGAGATTTGACGTCTGACCTTTGAGACCTTCTTCTCTTGG 2413
QY 1358 GTTTTCAGTGGCGGAGGAGTGGAGTTCTCTGTTTCCAGAAAGTGCATATCGGGACC 1417
DB 2414 CGTACCAAGTGGCCAAAGGATGGAATTCCTTGGAGTTCAAGTCTGTGTCCACAGAGACC 2473
QY 1418 TGGCAGCAGAAACATTTTATCTGAGAAACAGCTGGTGAAGATTTGCAATTTGGCC 1477
DB 2474 TGGCAGCAGAGATGTTGGTCTACCCAGGAGGTGGTGAAGATCTGTGACHTTGGAC 2533
QY 1478 TTGCCCGGATATTTAAGAACCCCGGATTAATGTGAGAAAGAGAGTACTGCTTCTC 1537
DB 2534 TGGCCCGAGACATCTGAGCGACTCCAGTCTGCTGAGGCGCAAGCAGCGGTGGCGG 2593
QY 1538 TGAATGATGATGCTCGGATCTATCTTTGACAAATCTACACGACCAAGAGAGAGCTGT 1597
DB 2594 TGAAGTGTGGCCCGGAGAGCTTTATTTGAAGGATCTTACAAATCAAGATGACGCT 2653
QY 1598 GGTCTTACGAGTATGCTGTGGAAATCTTCTCTTGTAGGTGGTCTCCATACCCAGAG 1657
DB 2654 GGTCTTACGAGTATGCTGTGGAGATTTTCTGCTGGGTGAGACCTTCTTCTGGA 2713
QY 1658 TCAATGATGAGGAGCTTTTGCAGTCCCTTCAAGGAGGAGCATGAGATGAGCTCTG 1717
DB 2714 TTTCTGTGCGGCTTACCTTATTAAGTGAATGATGAGTGTGATTTAAATGAGAGAGCAT 2773
QY 1718 AGTCTCTACTCTGGAATCTATGATCATCTGCTGGACTCTGCGACAGAGACCTCAAG 1777

DB 2774 TCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGCTTTTGACTCAGGA 2833
QY 1778 AAGGCGCAAGATTTGCAAGAACTTTGGAAACAACTAGTGATTTGCTTCAAGCAATGTAC 1837
DB 2834 ACGGCGCAATCTTCCCACTGACTTCAATTTAGGATGTCAGCTGGCGAGGAGCAAG 2893
QY 1838 AACAGGATGTTAAAGACTATCATCC 1861
DB 2894 AAGCATGATCAGAACATCATCC 2917

RESULT 14

US-09-872-136-1
Sequence 1, Application US/09872136
Patent No. US20020119545A1
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOIENT HEMATOPOIETIC STEM CELL
RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/872.136
FILING DATE: 01-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/208.786
FILING DATE: <Unknown>
APPLICATION NUMBER: US/09/021.324
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/977.451
FILING DATE: 1992-11-19
APPLICATION NUMBER: US 07/906.397
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: US PCT/US92/05401
FILING DATE: 26-JUN-1992
APPLICATION NUMBER: TW 81102961
FILING DATE: 15-APR-1992
APPLICATION NUMBER: US PCT/US92/02750
FILING DATE: 02-APR-1992
APPLICATION NUMBER: US 07/813.593
FILING DATE: 24-DEC-1991
APPLICATION NUMBER: US 07/793.065
FILING DATE: 15-NOV-1991
APPLICATION NUMBER: US 07/728.913
FILING DATE: 28-JUN-1991
APPLICATION NUMBER: US 07/679.666
FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-7P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3453 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 112..3006
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 31..111
FEATURE:
NAME/KEY: CDS
LOCATION: 31..3009
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-872-136-1

Query Match 10.5%: Score 221.6; DB 10; Length 3453;
Best Local Similarity 53.0%; Pred. No. 3.3e-55;
Matches 585; Conservative 0; Mismatches 489; Indels 30; Gaps 4;
785 GGCCTCCCTTATGATGCCACCAAGTGGAGTTTGGCGGGAGAGCTTAACTGGCAAT 844
Db 1817 GGGACTATGATGACCTTATGCTGGAGTTTCCGAGAGAGACTTAGAGTTTGGGAGG 1876
QY 845 CACTTGGAGAGGSGCTTTTGGAAAAGTGGTTCAAGCATCAGCAITTTGGCATTAGAAAT 904
Db 1877 TCCTGGGTCTGGCGTTTGGGGGGTGTATGACCCAGCGCTATGCAITATGATAAA 1936
QY 905 CACCTAGTCCCGAGCTGTGGCTGTGAATGTGAAGAGGGGCGCCAGCGAGGT 964
Db 1937 CGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAAGCTGACAGCTGTGAAA 1996
QY 965 ACAAGCTCTGATGACTGAGCTTAAATCTTGACCCACATTTGCCACCATCAGCAITTTGGCATTAGAAAT 904
Db 1977 AAGAGTCTCATGTCGGAGCTAAATGATGACCCAGCGCTATGCAITATGATAAA 1936
QY 1025 TTAACCTGTGGAGCTGTGACCAAGAGAGGGGCTCTGATGTGATTTGATTAAT 1084
Db 2057 TGAATCTGTGGGGGATGAC--ACTGTGAGGGGCTGATTTGATTTGATTAAT 2113
QY 1085 GCAATATGAATCTCTCACTACCTCAAGAGCAAGCTG----- 1126
Db 2114 GTTCTATGTTGACCTCTCACTACCTCAAGAGCAAGCTG----- 1126
QY 1127 ----ACTTATTTTTCACCAAGAGTCCAGCACTACATGGAGCTTAAGAAAGAAA 1182
Db 2174 GGACAGAGATTTTAAAGAACATAATTTCAGTTTACCTTACCTTACCTTACCTTAC 2233
QY 1183 ATGAGCCAGGCTGTGACAGAGGCAAGAACCAAGACTAGATAGGTACCCAGCGAA 1242
Db 2234 ATCCAGATGCTGTGTTTACAGAGAGTTCAGTTACACCGCGCTTGGATCAGCTCAG 2293
QY 1243 AG--CTTTGGAGCTCCGCTTTTCCAGAGATTAAGAGTGTGATGTTTGGAGAA-- 1298
Db 2294 GTTCAATGGAATCAATTTCTGAGATGAGATGAATATGAACCAAGAGAGG 2353
QY 1299 -GGAGATTTGACGTTTTCACAGAGGCGCCATCATTGGAAGATCTGATTTTATCA 1357
Db 2354 TGGCAGAGAGAGAGGAGAGATTTGAAGCTGTGAGCTTTGAGACCTCTTTGCTT 2413
QY 1358 GTTTCAAGTGGCAGAGGATGAGTTTCTGTTTCCAGAAAGTGCATTCAGGAGCC 1417
Db 2414 CGTACCAAGTGGCCAAAGGAGTGAATTTCTGGAGTTCAAGTCTGTGTGTCACAGAG 2473
QY 1418 TGGCAGGAGAGATCTTTTATCTGAGAACACGTTGGTGAAGTTTGTATTTTGGCC 1477
Db 2474 TGGCAGGAGAGATTTTGTGTCACCCAGGAGGTTGGTGAAGATCTGTGATTTTGGAC 2533
QY 1478 TTGCCCGGAGATTTTATGAAGACCCGATTTATGTGAGAAAGGATATCTGCTTCC 1537
Db 2534 TGGCCCGGAGATCTTGTAGGAGCTCCAGCTACCTGTGTCAGGGGCAAGCAGCGCT 2593
QY 1538 TGAATGGATGCTCCCAATCTATCTTTGACAAAATCTACAGCAACCAAGAGGCGGT 1597

Db 2594 TGAAGTGCATGGCAGCCGAGAGCTTATTTGAAGGATCTACACAATCAAGAGTGCCT 2653
QY 1598 GGTCTTACGGAGTATGCTGTGGAAATCTTCTCTTCTAGTGGTCTCCATACCCAGAG 1657
Db 2654 GGTCTTACGGAGTATGCTGTGGAGATATTTTCACTGGGTGTGAACCTTACCTTGGCA 2713
QY 1658 TACAATGGATGAGGAGCTTTTGCAGTGCCTTGAAGGAGGAGGATGAGAGCTCTG 1717
Db 2714 TTTCTGTGAGGCTTAACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAGAGCA 2773
QY 1718 AGTACTGTACTCTGAAATCTATCATGATCTGTGGAGTGTGGCAGAGAGAGCAAAAG 1777
Db 2774 TCTATGCCAGAGAGGATATCTTTTGAATGCAATCTCTCTGGGCTTTTGAATCAAGGA 2833
QY 1778 AAGGCCAAGATTTGCAAGACTTTGTGAAAACACTAGTGATTTGTTCAAGCAATATGAC 1837
Db 2834 AGCGCCATCTTCCCAACCTGACCTTCACTTTTATGAGTGTAGCTGGCAGGAGCAAG 2893
QY 1838 AACAGGATGTTAAAGACTACATCC 1861
Db 2894 AAGCATGTATCAGAACATCCATCC 2917

RESULT 15

US-09-866-510-1
: Sequence 1, Application US/09866510
: Patent No. US20020111304A1
: GENERAL INFORMATION:
: APPLICANT: KAZLAUSKAS, ANDRIUS
: APPLICANT: IKUNO, YASUSHI
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
: FILE REFERENCE: ERM-104.01
: CURRENT APPLICATION NUMBER: US/09/866,510
: PRIOR FILING DATE: 2001-05-25
: PRIOR APPLICATION NUMBER: 60/250,747
: PRIOR FILING DATE: 2000-12-01
: PRIOR APPLICATION NUMBER: 60/289,103
: PRIOR FILING DATE: 2001-05-07
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 1
: LENGTH: 3270
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(3267)
US-09-866-510-1

Query Match 10.2%: Score 213.4; DB 10; Length 3270;
Best Local Similarity 51.6%; Pred. No. 8.7e-53;
Matches 521; Conservative 0; Mismatches 486; Indels 96; Gaps 2;
QY 744 AATGACCCAGATGAAGTTCTTTGGATGAGCAGTGTGAGCGGCTCCCTTATGATGCCAG 803
Db 1592 AATCAGCCCGATGAGCATGATATTTATGTTGGACCGCTGAGCTGCTTATGACTC 1751
QY 804 CAAGTGGAGTTTCCCGGAGAGACTTAAATCATTGGCAATCATTGGAAGAGGCTTT 863
Db 1752 AAGATGGAGTTTCCAAAGAGATGAGCTAGTGTGTTGGGCTTGGGCTGAGAGCT 1811
QY 864 TGGAAAGTGGTTCAGCATCAGCATTTGCAATTTAGAAATCACCCTAGTGTGCGGACTGT 923
Db 1812 TGGAGAGTGGTTCAGCAAGAGAGAGCTTATGATTAAGCCCGTCCCACTGTGATGAAGT 1871
QY 924 GGCCTGTGAAATGCTGAAAGAGGCGCCAGCGGAGTACAAAGCTCTGATGACTCA 983
Db 1872 TGCAGTGAAGTGTATTAACCCAGCGGAGCTCCAGTGAAMACAGCTCTGATGCTGA 1931
QY 984 GCTAAATTTTGAACCCACATTTGCCACCATCTGAAAGCTGTGTTAACTGTGGAGCTG 1043
Db 1932 ACTGAAGATAATGACTACCTTGGGGGCCACATTTGAACATTTGTAACATTTGCTGGAGCTG 1991

This clone (DKFZp34L0121) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

location/Qualifiers
1..666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp34L0121"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site:1: NotI; Site:2: SalI"
BASE COUNT 195 a 146 c 157 g 164 t 4 others
ORIGIN

Query Match 26.7%; Score 561.4; DB 9; Length 666;
Best Local Similarity 99.8%; Pred. No. 3.4e-148;
Matches 562; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

/ 1539 GAATGAGTGGCTCCGAAATCTATCTTCGACAAATCTACAGACCAAGAGGACGAGTGG 1538
Dc 1 GAAATGGATGGCTCTGAAATCTATCTTCGACAAATCTACAGACCAAGAGGACGAGTGG 60
Qy 1599 GTCTACGAGGATTTGGTGGGAAATCTTCCTTAGGTGGGTCTCCATACCCAGGAGT 1638
Dc 61 GCTTACGAGGATTTGGTGGGAAATCTTCCTTAGGTGGGTCTCCATACCCAGGAGT 120
Qy 1659 ACAATGGATGAGGACTTTTCAGTGGCTCGCTGAGGAGGCAAGGATGAGGATCCCTGA 1718
Dc 121 ACAATGGATGAGGACTTTTCAGTGGCTCGCTGAGGAGGCAAGGATGAGGATCCCTGA 180
Qy 1719 GACCTCTCTCTGAAATCTATCAGATCATGCTGGACTGCTGGCAGAGACGACCAAGA 1778
Dc 181 GACCTCTCTCTGAAATCTATCAGATCATGCTGGACTGCTGGCAGAGACGACCAAGA 240
Qy 1779 AAGCCCAAGATTTGCAAACTTTGGGAAATCTAGTGTATTTCTTCAAGCAAAATGACA 1836
Dc 241 AAGCCCAAGATTTGCAAACTTTGGGAAATCTAGTGTATTTCTTCAAGCAAAATGACA 300
Qy 1939 ACAGATGTTAAAGACTACATCCATCAATCCATCAATCCATCAATCAATCAATCAAT 1896
Dc 301 ACAGATGTTAAAGACTACATCCATCAATCCATCAATCCATCAATCAATCAATCAAT 360
Qy 1899 ATACTCAACTCTGCTCTCTCTGAGGACTCTCTCAAGGAAAGTATTTTCAGTCCGAAGT 1958
Dc 361 ATACTCAACTCTGCTCTCTCTGAGGACTCTCTCAAGGAAAGTATTTTCAGTCCGAAGT 420
Qy 1959 TAATCAGGAGTCTGATGATGTCAGATATGTAATGCTTTCAAGTTCAGGAGCTGGA 2018
Dc 421 TAATCAGGAGTCTGATGATGTCAGATATGTAATGCTTTCAAGTTCAGGAGCTGGA 480
Qy 2019 AAGAAATCAAACTTTGAGAACTTTTACCGAATCCCACTCCCACTTCAGTTCAGTACCA 2078
Dc 481 AAGAAATCAAACTTTGAGAACTTTTACCGAATCCCACTCCCACTTCAGTTCAGTACCA 540
Qy 2079 GGGCGACAGCAGCTCTGTGG 2101
Dc 541 GGGCGACAGCAGCTCTGTGG 563

```

RESULT 5
P3005458
LOCUS
DEFINITION
MR3-GN0186-241100-008-c10 GN0186 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG005458
VERSION
BG005458.1 GI:12447642
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 650)

AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brantani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE
Shotgun sequencing of the human transcriptome with cDNA expressed
sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
20020263

COMMENT:
Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Frudente 109, 4 andar, 01505-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICE Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3st2-MR3-GN0186-

241100-008-c10st3-2000-11-24st4=1)

Seq primer: puc 18 forward

High quality sequence stop: 594.

Location/Qualifiers

1..650

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="GN0186"

/dev_stage="Adult"

/note="Organ: placenta; normal; Vector: puc18; Site:1: SmaI

products derived from ORESIES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the puc 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

BASE COUNT 195 a 146 c 166 g 138 t 2 others

ORIGIN

Query Match 26.6%; Score 558.2; DB 12; Length 650;

Best Local Similarity 96.6%; Pred. No. 2.8e-147;

Matches 569; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 261 GGAAGAAATCTCCAGAAAGAAATTCAGATCAGAGATCAGAGACCAATACCTCC 340

Dc 25 GGAAGAAATCTCCAGAAAGAAATTCAGATCAGAGATCAGAGACCAATACCTCC 84

Qy 341 TGGAAACCTCAGTATCACACAGTGGCCATCAGGAGTCCACCATTTAGACTGTCATG 400

Dc 85 GGGAAACCTCAGGATCACACAGGAGGAGTCCACCATTTAGACTGTCATG 144

Qy 401 CTATGGTGTCCCGAGCCCTCAGTACATCTGGTTTAAAGAAACCAACCAATACCAAG 460

Dc 145 CTATGGTGTCCCGAGCCCTCAGTACATCTGGTTTAAAGAAACCAACCAATACCAAG 204

Qy 461 AGCTGGAATATTTAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 520

Dc 205 AGCTGGAATATTTAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 264

Qy 521 AGCATGAGGTGTCTATCTACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580

Dc 265 GCAATGAGGTGTCTATCTACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 324

Qy 581 CATACCTCAGTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 640

Dc 325 CATACCTCAGTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 384

Qy 641 GCACCTGTGTGGCTCGGACTCTCTCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 700

Dc 365 GCACCTGTGTGGCTCGGACTCTCTCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 444

Qy 701 AAGGTCTTCTTCTGAAATAAGAGACTGACTTACTTACTTACTTACTTACTTACTTACTT 760


```

Db 445 AAGGTCCTCTTCGGAAATTAAGACTGACTACCTATCAATATATAATGAGCCAGAGAG 504
Qy 761 TTCTTTGGATGAGCAGTGTGCGGCTCCCTTATGATGCGCAAGAGTGGAGTTTGGCC 820
Db 505 TTCTTTGGATGAGCAGTGTGCGGCTCCCTTATGATGCGCAAGAGTGGAGTTTGGCC 564
Qy 821 GGGAGAGACTTAACTGGGCAATCACTTGGNAGAGGGGCTTTTGGAAA 869
Db 565 GGGAGAGACTTAACTGGGCAATCACTTGGNAGAGGGGCTTTTGGAAA 613

RESULT 6
AL575294/c 1000 bp mRNA linear EST 16-FEB-2001
LOCUS AL575294 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1061Y018 3
prime, mRNA sequence.
ACCESSION AL575294
VERSION AL575294.1 GI:12936322
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Li W.B., Gruber C., Jesse J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
source 1..1000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1061Y018"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
SE COUNT 242 a 199 c 230 g 320 t 9 others
IGIN

Query Match 22.2%; Score 466.8; DB 9; Length 1000;
Best Local Similarity 99.6%; Pred. No. 3.6e-121;
Matches 468; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TATATCAGAGTGTCCCAATGGGTTTCATGTTAACTTGGAAAAATGCGGAGGAGGA 60
Db 934 TATATCAGAGTGTCCCAATGGGTTTCATGTTAACTTGGAAAAATGCGGAGGAGGA 875
Qy 61 GAGGACCTCAACTCTCTTCAGAGTGTACAGTCTTATACAGAGAGTTCAGT 120
Db 874 GAGGACCTCAACTCTCTTCAGAGTGTACAGTCTTATACAGAGAGTTCAGT 815
Qy 121 TTACTCGGACAGTTTAAATACAGAAATGCGACATACAGTATTACAGAAATATGCCC 180
Db 814 TTACTCGGACAGTTTAAATACAGAAATGCGACATACAGTATTACAGAAATATGCCC 755
Qy 181 ATACATAGGAGCACTCCATCCTCTTAACTTACCTTCATGAAATGTTCCCTGCAAGAT 240
Db 754 ATACATAGGAGCACTCCATCCTCTTAACTTACCTTCATGAAATGTTCCCTGCAAGAT 695
Qy 241 TCAGGACACCTATGCTTCGAGAGCCAGGAATGTATACAGGGGAGAAATCTCTCCAGAG 300

```

```

Db 694 TCAGGACACCTATGCTTCGAGAGCCAGGAATGTATACAGGGGAGAAATCTCTCCAGAG 635
Qy 301 AAAGAAATTAACAATCAGATCAGGAGCAGCATACCTCTCGGAAACCTCAGTCAATCAC 360
Db 634 AAAGAAATTAACAATCAGATCAGGAGCAGCATACCTCTCGGAAACCTCAGTCAATCAC 575
Qy 361 AGATGGGCATCAGCAGTTCCACCATTTAGACTGTGATGCTAATGGTGTCCCGAGGCT 420
Db 574 AGATGGGCATCAGCAGTTCCACCATTTAGACTGTGATGCTAATGGTGTCCCGAGGCT 515
Qy 421 CAGATCACTTGGTITTAATAAACACACCAACCAATACACAGAGGCTGGAT 470
Db 514 CAGATCACTTGGTITTAATAAACACACCAACCAATACACAGAGGCTGGAT 465

RESULT 7
BQ770907 566 bp mRNA linear EST 26-JUL-2002
LOCUS BQ770907
DEFINITION IMAGE:6400365 5', mRNA sequence.
ACCESSION BQ770907
VERSION BQ770907.1 GI:21979381
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 566)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@pshs-femail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
source 1..566
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6400365"
/clone_lib="NIH_BMAP_F10"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (11 phage resistant)"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a NotI site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the NotI site and the polyA tail
is CAGCCAGCAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
Developing Mouse Nervous System, supported by National
Institute of Mental Health (NIMH), Hsien Chin, Ph.D.,
program coordinator."
BASE COUNT 149 a 146 c 154 g 116 t 1 others
IGIN

Query Match 21.5%; Score 451.6; DB 14; Length 566;

```


Db 719 TCTCCTATCGGAGAAATGTGGTTAAGATCTGTGACTTCGGGTGGCCCGGACCAITAT 778
QY 1494 TAAGAACCCCGATTATGTGAGAAAAGGAGATCTCGACTTCCTCTGAAATGGATGG 1549
Db 779 AAGA--CCGATTATGTACAGAAAGGAGATGCCCGAATCCCTTTGAGAGGTAIGG 832

Search completed: December 8, 2002, 12:49:57
Job time : 2741.79 secs